Christoph Schiller

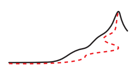
MOTION MOUNTAIN

THE ADVENTURE OF PHYSICS – VOL.VI

THE STRAND MODEL –

A SPECULATION ON UNIFICATION

www.motionmountain.net



Christoph Schiller

Motion Mountain

The Adventure of Physics
Volume VI

The Strand Model –
A Speculation on Unification

Edition 31, available as free pdf with films
at www.motionmountain.net/research.html

Editio trigesima prima.

Thirty-first edition.

To Britta, Esther and Justus Aaron

τῷ ἐμοὶ δαὶμονι

Man sollte manchmal einen kühnen Gedanken auszusprechen wagen,
damit er Frucht brächte.

PREFACE

This book is for anybody who is intensely curious about motion. Why and how do things, people, trees, stars, images or empty space move? The answer leads to many adventures, and this book presents one of the best of them: the search for a precise, unified and complete description of *all* motion.

The aim to describe *all* motion – everyday, quantum and relativistic – implies a large project. This project can be structured using the diagram shown in Figure 1, the so-called *Bronshtein cube*. The previous volumes have covered all points in the cube – all domains of motion – except the highest one. This remaining point contains the *complete* and *unified* description of all motion. The present volume briefly summarizes the history of this old quest and then presents an intriguing, though speculative solution to the riddle.

The search for the complete, unified description of motion is a story of many surprises. First, twentieth-century research has shown that there is a smallest measurable distance in nature, the Planck length. Then it appeared that matter cannot be distinguished from empty space at those small distances. A last surprise dates from this century: particles and space appear to be made of *strands*, instead of little spheres or points. The present text explains how to reach these surprising conclusions. In particular, quantum field theory, the standard model of particle physics, general relativity and cosmology are shown to follow from strands. The three gauge interactions, the three particle generations and the three dimensions of space turn out to be due to strands. In fact, all the open questions of twentieth-century physics about the foundations of motion, including the *origin of colours* and of the *constants of the standard model*, appear to be answerable.

The strand conjecture presented in this text is an unexpected result from a threefold aim that the author has pursued in the five previous volumes of this series: to present the basics of motion in a way that is up to date, captivating and simple. While the previous volumes introduced the *established* parts of physics, this volume presents, in the same captivating and playful way, a *speculation* about unification. Nothing in this volume is established knowledge – yet. The text is the original presentation of the topic. The aim for maximum simplicity has been central in deducing this speculation.

The search for a complete theory of motion is one of the adventures of life: it leads to the limits of thought. The journey overthrows several of our thinking habits about nature. This can produce fear, but by overcoming it we gain strength and serenity. Changing thinking habits requires courage, but it produces intense and beautiful emotions. Enjoy them.

Christoph Schiller

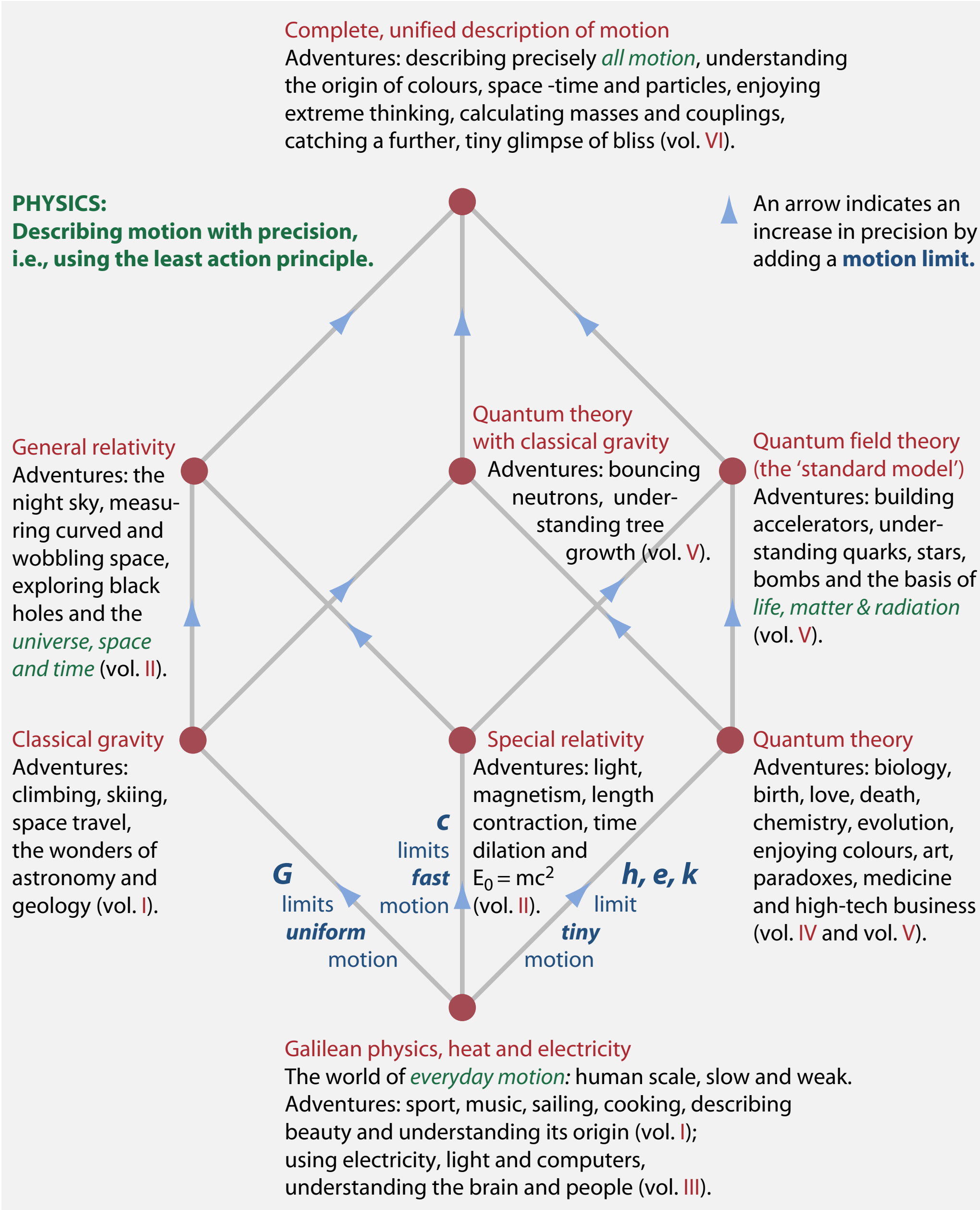


FIGURE 1 A complete map of physics, the science of motion, as first proposed by Matvei Bronshtein (b. 1907 Vinnytsia, d. 1938 Leningrad). The map is of central importance in the present volume. The Bronshtein cube starts at the bottom with everyday motion, and shows the connections to the fields of modern physics. Each connection increases the precision of the description and is due to a limit to motion that is taken into account. The limits are given for uniform motion by the gravitational constant G, for fast motion by the speed of light c, and for tiny motion by the Planck constant h, the elementary charge e and the Boltzmann constant k.

Using this book

To get a *fast overview*, read the pictorial summary in Chapter 13: one page of text and less than twenty pictures.

For a *short overview*, read the presentation of the quest in Chapter 1 and then continue with the summary sections at the end of each chapter.

The full text expounds the train of thoughts that led to the strand model. It is written for physics enthusiasts who like daring thoughts. Throughout the text,

▷ Important steps are marked with a triangle.

Also some dead ends are mentioned. Care has been taken not to include statements that disagree with experiment.

Marginal notes refer to bibliographic references, to other pages or to challenge solutions. In the colour edition, such notes and also the pointers to footnotes and to other websites are typeset in green. In the free pdf edition of this book, available at www.motionmountain.net, all green pointers and links are clickable. The pdf edition also contains all films; they can be watched directly in Adobe Reader. Over time, links on the internet tend to disappear. Most links can be recovered via www.archive.org, which keeps a copy of old internet pages.

Challenges are included regularly. Solutions and hints are given in the appendix. Challenges are classified as easy (e), standard student level (s), difficult (d) and research level (r). Challenges for which hints or solutions have not yet been included are marked (ny).

A *paper edition* of this book is available, either in colour or in black and white, from www.amazon.com. So is a Kindle edition.

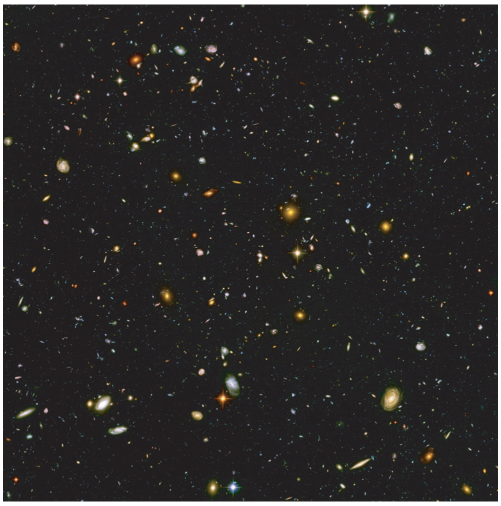
Feedback

Receiving an email from you at fb@motionmountain.net, either on how to improve the text or on a solution for one of the prize challenges mentioned on www.motionmountain.net/prizes.html, would be delightful. All feedback will be used to improve the next edition. For a particularly useful contribution you will be mentioned – if you want – in the acknowledgements, receive a reward, or both.

Support

Your donation to the minuscule, charitable, tax-exempt non-profit organisation that produces, translates and publishes this book series is welcome. For details, see the web page www.motionmountain.net/donation.html. The German tax office checks the proper use of your donation. If you want, your name will be included in the sponsor list. Thank you in advance for your help, on behalf of all readers across the world. And now, enjoy the reading.

CONTENTS



The Strand Model – A Speculation on Unification

Where, through the combination of
quantum theory and general relativity,
the top of Motion Mountain is reached,
and it is discovered
that vacuum is indistinguishable from matter,
that there is little difference between the large and the small,
that nature can be described by strands,
that particles can be modelled as tangles,
that gauge interactions appear naturally,
that colours are due to strand twisting,
that fundamental constants are uniquely fixed,
and that a complete description of motion is possible.

Chapter 1

FROM MILLENNIUM PHYSICS TO UNIFICATION

Look at what happens around us. A child that smiles, a nightingale that sings, a lily that opens: all move. Every shadow, even an immobile one, is due to moving light. Every mountain is kept in place by moving electrons. Every colour around us is due to the motion of electrons and photons. Every star owes its formation and its shine to motion of matter and radiation. Also the darkness of the night sky** is due to motion: it results from the expansion of space. Finally, human creativity and human actions are due to the motion of molecules, ions and electrons in the brain and in the body. Is there a common language for these and all other observations of nature?

Is there a *unified* and *precise* way to describe all motion? How? Is everything that moves, from people to planets, from light to empty space, made of the same constituents? What is the origin of motion? Answering these questions is the topic of the present text.

Answering questions about motion with precision defines the subject of *physics*. Over the centuries, researchers collected a huge number of precise observations about motion. We now know how electric signals move in the brain, how insects fly, why colours vary, how the stars formed, how life evolved, how the moon and the planets move, and much more. We use our knowledge about motion to look into the human body and heal illnesses; we use our knowledge about motion to build electronics, communicate over large distances, and work for peace; we use our knowledge about motion to secure life against many of nature's dangers, including droughts and storms. Physics, *the science of motion*, has shown time after time that knowledge about motion is both useful and fascinating.

At the end of the last millennium, humans were able to describe *all* motion in nature with high precision. This description can be summarized in the following six statements.

1. In nature, motion takes place in three dimensions of space and is described by the least action principle. Action is a physical quantity that describes how much change occurs in a process. The least action principle states: *motion minimizes change*. Among others, the least change principle implies that motion is predictable, that energy is conserved and that growth and evolution are natural processes, as is observed. Ref. 1, Ref. 3

2. In nature, there is an invariant maximum energy speed, the speed of light c. This invariant maximum implies *special relativity*. Among others, it implies that mass and energy are equivalent, as is observed. Ref. 2

** The photograph on page 16 shows an extremely distant, thus extremely young, part of the universe, with its large number of galaxies in front of the black night sky (courtesy NASA).

3. In nature, there is an invariant highest momentum flow, the Planck force $c^4/4G$. This invariant maximum implies *general relativity*, as we will recall below. Among others, general relativity implies that things fall and that empty space curves and moves, as is observed.
Page 31
Ref. 2

4. The evolution of the universe appears to be described by the cosmological constant Λ. Together with the largest distance and the largest age that can presently be observed, the cosmological constant determines the acceleration of the most distant stars.
Ref. 2

5. In nature, there is a non-zero, invariant smallest change value, the quantum of action $\hbar$. This invariant value implies *quantum theory*. Among others, it explains what life and death are, why they exist and how we enjoy the world.
Ref. 4

6. In nature, matter and radiation consist of quantum particles. Matter consists of *fermions*: six quarks, three charged leptons, three neutrinos and their antiparticles. Radiation consists of *bosons*: the photon, three intermediate weak vector bosons and eight gluons. In addition, the year 2012 finally brought the discovery of the Higgs boson, which was already predicted in 1964. Fermions and bosons move and can transform into each other. The transformations are described by the electromagnetic interaction, the weak nuclear interaction and the strong nuclear interaction. Together with the masses, quantum numbers, mixing angles and couplings of the elementary particles, these transformation rules form the so-called *standard model of particle physics*. Among others, the standard model explains how lightning forms, why colours vary, and how the atoms in our bodies came to be.
Ref. 4

These six statements, the *millennium description of physics*, describe everything known about motion in the year 2000. (Actually, 2012 is a more precise, though less striking date.) These statements describe the motion of people, animals, plants, objects, light, radiation, stars, empty space and the universe. The six statements describe motion so precisely that even today there is *no* difference between calculation and observation, between theory and practice. This is an almost incredible result, the summary of the efforts of tens of thousands of researchers during the past centuries.

However, a small set of observations does not yet follow from the six statements. A famous example is the origin of colours. In nature, colours are consequences of the so-called *fine structure constant*, a mysterious constant of nature, abbreviated α, whose value is measured to be $\alpha = 1/137.035\,999\,139(31)$. If α had another value, all colours would differ. And why are there *three* gauge interactions, *twelve* elementary fermions, *thirteen* elementary bosons and *three* dimensions? What is the origin of particle masses? Why is the standard model, the sixth statement above, so complicated? How is it related to the five preceding statements?
Ref. 5

A further unexplained observation is the nature of dark matter found around galaxies. We do not know yet what it is. Another unexplained process is the way thinking forms in our brain. We do not know yet in detail how thinking follows from the above six statements, though we do know that thinking is not in contrast with them. For this reason, we will not explore the issue in the following. In the case of dark matter this is not so clear: dark matter could even be in contrast with the millennium description of motion.

Finally, why is there motion anyway? In short, even though the millennium description of physics is precise and successful, it is not complete. The list of all those fundamental issues about motion that are *unexplained* since the year 2000 make up only a short table. We call them the *millennium issues*.

TABLE 1 The millennium list: *everything* the standard model and general relativity *cannot* explain; thus, also the list of the *only* experimental data available to test the complete description of motion.

Observable	Property unexplained since the year 2000
Local quantities unexplained by the standard model: particle properties	
$\alpha = 1/137.036(1)$	the low energy value of the electromagnetic coupling or fine structure constant
α_w or θ_w	the low energy value of the weak coupling constant or the value of the weak mixing angle
α_s	the value of the strong coupling constant at one specific energy value
m_q	the values of the 6 quark masses
m_l	the values of 6 lepton masses
m_W	the value of the mass of the W vector boson
m_H	the value of the mass of the scalar Higgs boson
$\theta_{12}, \theta_{13}, \theta_{23}$	the value of the three quark mixing angles
δ	the value of the CP violating phase for quarks
$\theta^{\nu}_{12}, \theta^{\nu}_{13}, \theta^{\nu}_{23}$	the value of the three neutrino mixing angles
$\delta^{\nu}, \alpha_1, \alpha_2$	the value of the three CP violating phases for neutrinos
$3 \cdot 4$	the number of fermion generations and of particles in each generation
J, P, C, etc.	the origin of all quantum numbers of each fermion and each boson
Concepts unexplained by the standard model	
$c, \hbar, k$	the origin of the invariant Planck units of quantum field theory
$3 + 1$	the number of dimensions of physical space and time
SO(3,1)	the origin of Poincaré symmetry, i.e., of spin, position, energy, momentum
Ψ	the origin and nature of wave functions
$S(n)$	the origin of particle identity, i.e., of permutation symmetry
Gauge symmetry	the origin of the gauge groups, in particular:
U(1)	the origin of the electromagnetic gauge group, i.e., of the quantization of electric charge, of the vanishing of magnetic charge, and of minimal coupling
SU(2)	the origin of weak interaction gauge group, its breaking and P violation
SU(3)	the origin of strong interaction gauge group and its CP conservation
Renorm. group	the origin of renormalization properties
$\delta W = 0$	the origin of the least action principle in quantum theory
$W = \int L_{SM}\, dt$	the origin of the Lagrangian of the standard model of particle physics
Global quantities unexplained by general relativity and cosmology	
0	the observed flatness, i.e., vanishing curvature, of the universe
$1.2(1) \cdot 10^{26}$ m	the distance of the horizon, i.e., the 'size' of the universe (if it makes sense)
$\rho_{de} = \Lambda c^4/(8\pi G) \approx 0.5\ \mathrm{nJ/m^3}$	the value and nature of the observed vacuum energy density, dark energy or cosmological constant

TABLE 1 (Continued) The millennium list: *everything* the standard model and general relativity *cannot* explain; also the *only* experimental data available to test the complete description of motion.

Observable	Property unexplained since the year 2000
$(5 \pm 4) \cdot 10^{79}$	the number of baryons in the universe (if it makes sense), i.e., the average visible matter density in the universe
ρ_{dm}	the density and nature of dark matter
$f_0(1, ..., c.\ 10^{90})$	the initial conditions for *c.* 10^{90} particle fields in the universe (if or as long as they make sense), including the homogeneity and isotropy of matter distribution, and the density fluctuations at the origin of galaxies
Concepts unexplained by general relativity and cosmology	
c, G	the origin of the invariant Planck units of general relativity
$\mathrm{R} \times \mathrm{S}^3$	the observed topology of the universe
$G^{\mu\nu}$	the origin and nature of curvature, the metric and horizons
$\delta W = 0$	the origin of the least action principle in general relativity
$W = \int L_{GR} \mathrm{d}t$	the origin of the Lagrangian of general relativity

The millennium list contains *everything* that particle physics and general relativity *cannot* explain. In other words, the list contains *every issue* that was unexplained in the domain of fundamental motion in the year 2000. The list is short, but it is not empty. Every line in the millennium list asks for an explanation. The quest for unification – and the topic of this text – is the quest for these explanations. We can thus say that a *complete theory of motion* is a theory that eliminates the millennium list of open issues.

How to name the result of the quest

An number of expressions have been used for the result of the present quest. The term *theory of everything* is found in many media. However, the expression is pompous and wrong; the quest does not have this aim. Many problems of physics, science and life remain unsolved even if the quest presented here comes to an end. The tern is now used mainly by esoteric healers for their unsuccessful healing practices.

The term *final theory* has also been popular. In the meantime, it is mainly used in titles of mediocre books and films.

The term *world formula* has been invented by German journalists. It never became popular among physicists. It is now used for calculating the optimal way to park a car backwards.

The quest to eliminate the millennium list of open issues implies to find a *complete theory of motion*, i.e., a *complete description of motion*. These are adequate and correct expressions. Also the expressions *unified theory of motion* or *unified description of motion* are acceptable. They stress that a complete description must combine general relativity and the standard model of particle physics.

Against a complete theory

We know that a complete theory exists: it is the theory that describes how to calculate the fine structure constant $\alpha = 1/137.036(1)$. The theory does the same for about two dozen other constants, but α is the most famous one. In other terms, the complete theory is the theory that explains all colours found in nature.

Searching for a complete theory of motion is a fascinating journey. But not everybody agrees. A number of arguments are repeated again and again *against* the search for a unified theory. Reaching the complete theory and enjoying the adventure is only possible if these arguments are known – and then put gently aside.

— It is regularly claimed that a complete theory cannot exist because nature is infinite and mysteries will always remain. But this statement is wrong. First, nature is not infinite. Second, even if it were infinite, knowing and describing everything would still be possible. Third, even if knowing and describing everything would be impossible, and if mysteries would remain, a complete theory remains possible. A unified theory is *not* useful for *every* issue of everyday life, such as choosing your dish on a menu or your future profession. A complete theory is simply a full description of the *foundations* of motion: the unified theory just combines and explains particle physics and general relativity.

— It is sometimes argued that a complete theory cannot exist due to Gödel's incompleteness theorem or due to computational irreducibility. However, in such arguments, Page 170 both theorems are applied to domains were they are not valid. The reasoning is thus wrong.

— Some state that it is not clear whether a complete theory exists at all. But we all know from experience that this is wrong. The reason is simple: *We are able to talk about everything.* In other words, all of us already have a 'theory of everything', or a complete theory of nature. Also a physical theory is a way to talk about nature, and for the complete theory we only have to search for those concepts that enable us to talk about all of motion *with full precision*. Because we are just looking for a way to talk, we know that the unified theory exists. And searching for it is fascinating and exciting, as everybody busy with this adventure will confirm.

— Some claim that the search for a unified theory is a reductionist endeavour and cannot Ref. 6 lead to success, because reductionism is flawed. This claim is wrong on three counts. First, it is not clear whether the search is a reductionist endeavour, as will become Page 162 clear later on. Second, there is no evidence that reductionism is flawed. Third, even if it were, no reason not to pursue the quest would follow. The claim in fact invites to search with a larger scope than was done in the past decades – an advice that will turn out to be spot on.

— Some argue that searching for a unified theory makes no sense as long as the measurement problem of quantum theory is not solved, or consciousness is not understood, Ref. 7 or the origin of life is not understood. Now, the measurement problem is solved by Vol. IV, page 143 decoherence, and in order to combine particle physics with general relativity, understanding the details of consciousness or of the origin of life is not required. Neither is understanding child education required – though this can help.

— Some people claim that searching for a complete theory is a sign of foolishness or a sin of pride. Such defeatist or envious comments should simply be ignored. After all, the quest is just the search for the solution to a riddle.

Ref. 8 — Some believe that understanding the unified theory means to read the mind of god, or to think like god, or to be like god. This is false, as any expert on god will confirm. In fact, solving a riddle or reading a physics textbook does not transform people into gods. This is unfortunate, as such an effect would provide excellent advertising.

— Some fear that knowing the complete theory will yield immense power that harbours huge dangers of misuse. In short, some claim that knowing the complete theory will change people into devils. However, this fear is purely imaginary; it only describes the fantasies of the person that is talking. Indeed, the millennium description of physics is already quite near to the complete theory, and nothing to be afraid of has happened. Sadly, another great advertising opportunity evaporates. Ref. 9
— Some people object that various researchers in the past have thought to have found the unified theory, but were mistaken, and that many great minds tried to find such a theory, but had no success. That is true. Some failed because they lacked the necessary data for a successful search, others because they lost contact with reality, and still others because they were led astray by prejudices that limited their progress. We just have to avoid these mistakes.

These arguments show us that we can reach the complete theory of motion – which we symbolically place at the top of Motion Mountain – only if we are not burdened with ideological or emotional baggage. (We get rid of all baggage in the first six chapters of this volume.) The goal we have set thus requires *extreme thinking*, i.e., thinking up to the limits. After all, unification is the precise description of *all* motion, including its most extreme cases.

Therefore, unification is, first of all, a riddle. The search for unification is a pastime. Any pastime is best approached with the light-heartedness of playing. Life is short: we should play whenever we can. Ref. 10

What went wrong in the past

The twentieth century was the golden age of physics. Scholars searching for the unified theory explored candidates such as grand unified theories, supersymmetry and numerous other options. These candidates will be discussed later on; all were falsified by experiment. Despite a record number of physicists working on the problem, despite the availability of extensive experimental data, and despite several decades of research, no unified theory was found. Why? Vol. V, page 269

During the twentieth century, many successful descriptions of nature were deformed into dogmatic beliefs about unification. Here are the main examples, with some of their best known proponents:

— ‘Unification requires generalization of existing theories.’
— ‘Unification requires finding higher symmetries.’ (Werner Heisenberg)
— ‘Unification requires generalizing electroweak mixing to include the strong interaction.’ (Abdus Salam)
— ‘Unification requires extending the standard model of particle physics with supersymmetry.’ (Steven Weinberg)
— ‘Unification requires axiomatization.’ (David Hilbert)
— ‘Unification requires searching for beauty.’ (Paul Dirac)
— ‘Unification requires new quantum evolution equations.’ (Werner Heisenberg)

— ‘Unification requires new field equations of gravitation.’ (Albert Einstein)
— ‘Unification requires more dimensions of space.’ (Theodor Kaluza)
— ‘Unification requires topology change.’ (John Wheeler)
— ‘Unification is independent of Planck’s natural units.’
— ‘Unification requires using complicated mathematics and solving huge conceptual difficulties.’ (Edward Witten)
— ‘Unification is only for a selected few.’
— ‘Unification is extremely useful, important and valuable.’

All these beliefs appeared in the same way: first, some famous scholar – in fact, many more than those mentioned – explained the idea that guided his own discovery; then, he and most other researchers started to believe the guiding idea more than the discovery itself. The most explored beliefs were those propagated by Salam and Weinberg: they – unknowingly – set thousands of researchers on the wrong path for many decades. The most detrimental belief was that unification is complicated and difficult: it kept the smartest physicists from producing progress. In fact, all the mentioned beliefs can be seen as special cases of the first one. And like the first belief, also all the others are, as we will discover in the following, wrong.

An encouraging argument

Page 8 The Bronshtein cube in Figure 1 shows that physics started from the description of motion in everyday life. At the next level of precision, physics introduced the observed *limits* to motion and added the description of powerful, i.e., as uniform as possible motion (classical gravity), as fast as possible motion (special relativity), and as tiny as possible motion (quantum theory). At the following level of precision, physics achieved all possible combinations of two of these motion types, by taking care of two motion limits at the same time: fast and uniform motion (general relativity), fast and tiny motion (quantum field theory), and tiny and uniform motion (quantum theory with gravity). The only domain left over is the domain where motion is fast, tiny and as uniform as possible at the same time. When this last domain is reached, the precise description of all motion is completed.

But Figure 1 suggests even stronger statements. First of all, no domain of motion is left: the figure covers *all* motion. Secondly, the unified description appears when general relativity, quantum field theory and quantum theory with gravity are combined. In other words, the unified theory appears when relativity *and* quantum theory *and* interactions are described together. But a third conclusion is especially important. Each of these three fields of physics can be deduced from the unified theory by eliminating a limitation: either that of tiny motion, that of straight motion, or that of fast motion. In other words:

▷ General relativity follows from the unified theory by eliminating the quantum of action $\hbar$, i.e., taking the limit $\hbar \rightarrow 0$.

▷ Quantum field theory, including quantum electrodynamics, follows from the unified theory by eliminating G, i.e., taking the limit $G \rightarrow 0$.

▷ Quantum theory with gravity follows from the unified theory by eliminating the speed limit c, i.e., taking the limit $1/c \rightarrow 0$.

Speaking even more bluntly, and against a common conviction of researchers in the field, the figure suggests: *The standard model follows from the complete theory by eliminating gravity.* These connections eliminate many candidates for the unified theory that were proposed in the research literature in the twentieth and twenty-first century. But more importantly, the connections leave open a range of possibilities – and interestingly enough, this range is very narrow.

Figure 1 allows stronger statements still. Progress towards the unified theory is achieved by taking limitations to motion into account. Whatever path we take from everyday physics to the unified theory, we must take into consideration all limits to motion. The order can differ, but *all* limits have to be taken into account. Now, if any intermediate steps – due to additional motion limitations – between quantum field theory and the unified theory existed in the *upper* part of the figure, corresponding steps would have to appear also in the *lower* part of the figure, between everyday physics and classical gravity. Since no such intermediate steps exist, there are no intermediate steps between the standard model and the unified theory. In the same way, if any intermediate limits or steps between general relativity and the complete theory really existed, these limits and the corresponding steps would also have to appear between everyday motion and quantum theory.

Experiments show clearly that *no* intermediate steps or limits exist between everyday motion and the next level of precision. Using the top-down symmetry of Figure 1, this implies:

▷ Intermediate steps or theories do not exist before the complete theory.

This is a strong statement. In the foundations of motion, apart from the unified theory, no further theory is missing. For example, the Bronshtein cube implies that there is no separate theory of relativistic quantum gravity or no doubly special relativity.

Figure 1 also implies that, conceptually, we are already *close* to the unified theory. The figure suggests that there is no need for overly elaborate hypotheses or concepts to reach the complete theory:

▷ We just have to add G to the standard model or $\hbar$ and e to general relativity.

In short, the complete, unified theory of motion cannot be far.

Summary: how to find the complete theory of motion

We have a riddle to solve: we want to describe precisely all motion and discover its origin. In order to achieve this, we need to find a complete theory that solves and explains each open issue given in the *millennium list*. This is our starting point.

We proceed in steps. We first simplify quantum theory and gravitation as much as possible, we explore what happens when the two are combined, and we deduce the *requirement list* that any unified theory must fulfil. Then we deduce the simplest possible model that fulfils the requirements; we check the properties of the model against every experiment performed so far and against every open issue from the millennium list. Discovering that there are no disagreements, no points left open and no possible alternatives, we know that we have found the unified theory. We thus end our adventure with a *list of testable predictions* for the proposed theory.

In short, three lists structure our quest for a complete theory: the millennium list of open issues, the list of requirements for the complete theory, and the list of testable predictions. To get from one list to the next, we proceed along the following legs.

1. We first simplify modern physics. Twentieth century physics deduced several *invariant* properties of motion. These invariants, such as the speed of light or the quantum of action, are called *Planck units*. The invariant Planck units allow motion to be measured. Above all, these invariants are also found to be *limit values*, valid for every example of motion.
2. Combining quantum theory and general relativity, we discover that at the Planck limits, the universe, space and particles are *not described by points*. We find that as long as we use points to describe particles and space, and as long as we use sets and elements to describe nature, a unified description of motion is impossible.
3. The combination of quantum theory and general relativity teaches us that space and particles have *common constituents*.
4. By exploring black holes, spin, and the limits of quantum theory and gravity, we discover that the common constituents of space and particles are *extended*, without ends, one-dimensional and fluctuating: the common constituents of space and particles are *fluctuating strands*.
5. We discover that we cannot think or talk without continuity. We need a *background* to describe nature. We conclude that to talk about motion, we have to combine continuity and non-continuity in an appropriate way. This is achieved by imagining that fluctuating strands move in a continuous three-dimensional *background*.

Page 149 At this point, after the first half of our adventure, we obtain a detailed *requirement list* for the unified theory. This list allows us to proceed rapidly towards our goal, without being led astray:

6. We discover a simple fundamental principle that explains how the maximum speed c, the minimum action $\hbar$, the maximum force $c^4/4G$ and the cosmological constant Λ follow from strands. We also discover how to deduce quantum theory, relativity and cosmology from strands.
7. We discover that strands naturally yield the existence of three spatial dimensions, flat and curved space, black holes, the cosmological horizon, fermions and bosons. We find that all known physical systems are made from strands. Also the process of measurement and all properties of the background result from strands.
8. We discover that fermions emit and absorb bosons, and that they do so with exactly those properties that are observed for the electromagnetic, the weak and the strong nuclear interaction. In short, the *three known gauge interactions* – and their parity conservation or violation – follow from strands in a unique way. In addition, we discover that other interactions do not exist.
9. We discover that strands naturally yield the known elementary fermions and bosons,

grouped in *three generations*, and all their observed properties. Other elementary particles do not exist. We thus recover the standard model of elementary particles.

10. We discover that the fundamental principle allows us to solve all the issues in the millennium list, and that all properties deduced from strands agree with experiment. In particular, the strand conjecture allows us to calculate the fine structure constant and the other gauge coupling strengths. An extensive *list of testable predictions* arises. These predictions will all be tested – by experiment or by calculation – in the coming years.
Page 412
11. We discover that motion is due to crossing switches of strands. Motion is an inescapable consequence of observation: motion is an experience that we make because we are, like every observer, a small, approximate part of a large whole.

At the end of this journey, we will thus have unravelled the mystery of motion. It is a truly special adventure. *But be warned: almost all of the story presented here is still speculative, and thus open to question.* Everything presented in the following agrees with experiment. Nevertheless, with almost every sentence you will find at least one physicist or philosopher who disagrees. That makes the adventure even more fascinating.

> Es ist fast unmöglich, die Fackel der Wahrheit durch ein Gedränge zu tragen, ohne jemandem den Bart zu sengen.*
>
> Georg Christoph Lichtenberg

* 'It is almost impossible to carry the torch of truth through a crowd without scorching somebody's beard.' Georg Christoph Lichtenberg (b. 1742 Ober-Ramstadt, d. 1799 Göttingen) was a famous physicist and essayist.

Chapter 2

PHYSICS IN LIMIT STATEMENTS

Twentieth century physics deduced several *invariant* properties of motion. These invariants, such as the speed of light or the quantum of action, define the so-called Planck units. The invariant Planck units are important for two reasons: first, they allow motion to be measured; second, the invariants are *limit values*. In fact, the Planck units provide bounds for all observables.

The main lesson of modern physics is thus the following: When we simplify physics as much as possible, we discover that *nature limits the possibilities of motion*. Such limits lie at the origin of special relativity, of general relativity and of quantum theory. In fact, we will see that nature limits *every* aspect of motion. Exploring the limits of motion will allow us to deduce several astonishing conclusions. And these conclusions contradict all that we learned about nature so far.

SIMPLIFYING PHYSICS AS MUCH AS POSSIBLE

At dinner parties, physicists are regularly asked to summarize physics in a few sentences. It is useful to have a few simple statements ready to answer such a request. Such statements are not only useful to make other people think; they are also useful in our quest for the unified theory. Here they are.

Everyday, or Galilean, physics in one statement

Everyday motion is described by Galilean physics. It consists of only one statement:

▷ Motion minimizes change.

In nature, *change* is measured by physical action W. More precisely, change is measured by the time-averaged difference between kinetic energy T and potential energy U. In other words, all motion obeys the so-called *least action principle*. It can be written as

$$\delta W = 0 \text{ , where } \quad W = \int (T - U)\, \mathrm{d}t \ . \tag{1}$$

This statement determines the effort we need to move or throw stones, and explains why cars need petrol and people need food. In simpler terms, *nature is as lazy as possible*. Or:

▷ Nature is maximally efficient.

Vol. I, page 29 Inefficient motion is not observed. The efficiency and laziness of nature implies that motion is conserved, relative and predictable. In fact, the laziness of motion and nature is valid throughout nature. A few additional ways to distinguish observed motion from impossible motion were discovered by modern physics.

Special relativity in one statement

The step from everyday, or Galilean, physics to special relativity can be summarized in a Ref. 11 single limit statement on motion. It was popularized by Hendrik Antoon Lorentz:

▷ There is a maximum energy speed value c in nature.

For all physical systems and all observers, the local energy speed v is limited by the speed of light c:

$$v \leqslant c = 3.0 \cdot 10^8 \text{ m/s} . \qquad (2)$$

All results peculiar to special relativity follow from this principle. A few well-known facts set the framework for the discussion that follows. The speed v is less than or equal to the speed of light c for *all* physical systems;* in particular, this speed limit is valid both for composite systems and for elementary particles, for matter and radiation. No exception Challenge 1 e has ever been found. (Try it.)

The energy speed limit is an *invariant*: the local energy speed limit is valid for *all* observers. In this context it is essential to note that any observer must be a physical system, Vol. II, page 100 and must be *close* to the moving energy.

The speed limit c is realized by *massless* particles and systems; in particular, it is realized by electromagnetic waves. For matter systems, the speed is always below c.

Only a maximum energy speed ensures that cause and effect can be distinguished in nature, or that sequences of observations can be defined. The opposite hypothesis, that energy speeds greater than c are possible, which implies the existence of so-called *(real) tachyons*, has been explored and tested in great detail; it leads to numerous conflicts with observations. Tachyons do not exist.

The maximum energy speed forces us to use the concept of *space-time* to describe nature, because the existence of a maximum energy speed implies that space and time Vol. II, page 27 *mix*. It also implies observer-dependent time and space coordinates, length contraction, time dilation, mass–energy equivalence, horizons for accelerated observers, and all the other effects that characterize special relativity. Only a maximum speed leads to the principle of maximum ageing that governs special relativity; and only this principle leads to the principle of least action at low speeds. In addition, only with a finite speed limit is it possible to define a *unit* of speed that is valid at all places and at all times. If there

* A *physical system* is a region of space-time containing mass–energy, the location of which can be followed over time and which interacts incoherently with its environment. The speed of a physical system is thus an *energy speed*. The definition of physical system excludes images, geometrical points or incomplete, entangled situations.

were no global speed limit, there could be no natural measurement standard for speed, independent of all interactions; speed would not then be a measurable quantity.

Special relativity also limits the size of systems – whether composite or elementary. Indeed, the limit speed implies that acceleration a and size l cannot be increased independently without bounds, because the two ends of a system must not interpenetrate. The most important case concerns massive systems, for which we have

$$l \leqslant \frac{c^2}{a} . \tag{3}$$

This size limit is induced by the speed of light c; it is also valid for the *displacement* d of a system, if the acceleration measured by an external observer is used. Finally, the speed limit implies a relativistic 'indeterminacy relation'

$$\Delta l \, \Delta a \leqslant c^2 \tag{4}$$

for the length and acceleration indeterminacies. You may wish to take a minute to deduce this relation from the time–frequency indeterminacy. All this is standard knowledge.

Challenge 2 s

Quantum theory in one statement

The difference between Galilean physics and quantum theory can be summarized in a single statement on motion, due to Niels Bohr:

Ref. 12

▷ There is a minimum action value $\hbar$ in nature.

Vol. IV, page 15

For all physical systems and all observers, the action W obeys

$$W \geqslant \hbar = 1.1 \cdot 10^{-34}\,\mathrm{Js} . \tag{5}$$

The Planck constant $\hbar$ is the smallest observable action value, and the smallest observable change of angular momentum. The action limit is valid for all systems, thus both for composite and elementary systems. No exception has ever been found. (Try it.) The principle contains all of quantum theory. We can call it the *principle of non-zero action*, in order to avoid confusion with the principle of least action.

Challenge 3 e

The non-zero action limit $\hbar$ is an *invariant*: it is valid with the same numerical value for *all* observers. Again, any such observer must be a physical system.

The action limit is realized by many physical processes, from the absorption of light to the flip of a spin 1/2 particle. More precisely, the action limit is realized by *microscopic* systems where the process involves a single particle.

The non-zero action limit is stated less frequently than the speed limit. It starts from the usual definition of the action, $W = \int (T-U)\,\mathrm{d}t$, and states that between two observations performed at times t and $t + \Delta t$, even if the evolution of a system is not known, the measured action is at least $\hbar$. Since physical action measures the change in the state of a physical system, there is always a minimum change of state between two different observations of a system.* The non-zero action limit expresses the fundamental fuzziness of nature at a microscopic scale.

It can easily be checked that no observation – whether of photons, electrons or macroscopic systems – gives a smaller action than the value $\hbar$. The non-zero action limit has been verified for fermions, bosons, laser beams, matter systems, and for any combination of these. The opposite hypothesis, implying the existence of arbitrary small change, has been explored in detail: Einstein's long discussion with Bohr, for example, can be seen as a repeated attempt by Einstein to find experiments that would make it possible to measure arbitrarily small changes or action values in nature. In every case, Bohr found that this could not be achieved. All subsequent attempts were equally unsuccessful.

Ref. 13 The principle of non-zero action can be used to deduce the indeterminacy relation, the tunnelling effect, entanglement, permutation symmetry, the appearance of probabilities in quantum theory, the information-theoretic formulation of quantum theory, and the existence of elementary particle reactions. Whenever we try to overcome the smallest action value, the experimental outcome is probabilistic. The minimum action value also implies that in quantum theory, the three concepts of state, measurement operation, and measurement result need to be distinguished from each other; this is done by means Ref. 14 of a so-called *Hilbert space*. Finally, the non-zero action limit is also the foundation of Einstein–Brillouin–Keller quantization.

The existence of a non-zero action limit has been known from the very beginning of quantum theory. It is at the basis of – and completely equivalent to – all the usual formulations of quantum theory, including the many-path and the information-theoretic formulations.

We also note that only a non-zero action limit makes it possible to define a *unit* of action. If there were no action limit, there could be no natural measurement standard for action: action would not then be a measurable quantity.

The upper bounds for speed and for action for any physical system, $v \leqslant c$ and $W \leqslant pd \leqslant mcd$, when combined with the quantum of action, imply a limit on the displacement d of a system between any two observations:

$$d \geqslant \frac{\hbar}{mc} . \tag{6}$$

In other words, the (reduced) Compton wavelength of quantum theory appears as the lower limit on the displacement of a system, whenever gravity plays no role. Since this quantum displacement limit also applies to elementary systems, it also applies to the *size* of a *composite* system. However, for the same reason, this size limit is *not* valid for the Challenge 5 e sizes of elementary particles.

The limit on action also implies Heisenberg's well-known indeterminacy relation for Vol. IV, page 24 the displacement d and momentum p of physical systems:

$$\Delta d \, \Delta p \geqslant \frac{\hbar}{2} . \tag{7}$$

* For systems that seem constant in time, such as a spinning particle or a system showing the quantum Zeno Challenge 4 e effect, finding this minimum change is tricky. Enjoy the challenge.

This relation is valid for both massless and massive systems. All this is textbook knowledge.

Thermodynamics in one statement

Thermodynamics can also be summarized in a single statement about motion:

▷ There is a smallest entropy value k in nature.

Written symbolically,

$$S \geqslant k = 1.3 \cdot 10^{-23}\,\mathrm{J/K}\,. \tag{8}$$

Challenge 6 e The entropy S is limited by the Boltzmann constant k. No exception has ever been found. Ref. 15 (Try it.) This result is almost 100 years old; it was stated most clearly by Leo Szilard. All of thermodynamics can be deduced from this relation, in combination with the quantum of action.

The entropy limit is an *invariant*: it is valid for *all* observers. Again, any observer must be a physical system.

The entropy limit is realized only by physical systems made of a single particle. In other words, the entropy limit is again realized only by *microscopic* systems. Therefore the entropy limit provides the same length limit for physical systems as the action limit.

Like the other limit statements we have examined, the entropy limit can also be phrased as a indeterminacy relation between temperature T and energy U:

$$\Delta\frac{1}{T}\,\Delta U \geqslant \frac{k}{2}\,. \tag{9}$$

Ref. 16 This relation was first given by Bohr and then discussed by Heisenberg and many others.

General relativity in one statement

This text can be enjoyed most when a compact and unconventional description of general relativity is used; it is presented in the following. However, the conclusions do not depend on this description; the results are also valid if the usual approach to general relativity is Page 55 used; this will be shown later on.

The most compact description summarizes the step from universal gravity to general relativity in a single statement on motion:

▷ There are maximum force and power values in nature.

For all physical systems and all observers, force F and power P are limited by

$$F \leqslant \frac{c^4}{4G} = 3.0 \cdot 10^{43}\,\mathrm{N} \quad \text{and} \quad P \leqslant \frac{c^5}{4G} = 9.1 \cdot 10^{51}\,\mathrm{W}\,. \tag{10}$$

Challenge 7 e No exception has ever been found. (Try it.) These limit statements contain both the speed of light c and the gravitational constant G; they thus qualify as statements about relativistic gravitation. Before we deduce general relativity, let us explore these limits.

The numerical values of the limits are huge. The maximum power corresponds to radiating 50 solar masses within 1 millisecond. And applying the maximum force value along a distance l requires as much energy as is stored in a black hole of diameter l.

Force is change of momentum; power is change of energy. Since momentum and energy are conserved, force and power are the flow of momentum and energy *through a* Ref. 17 *surface*. Force and power, like electric current, describe the change in time of conserved quantity. For electric current, the conserved quantity is charge, for force, it is momentum, Vol. I, page 231 for power, it is energy. In other words, like current, also force is a flow across a surface. This is a simple consequence of the continuity equation. Therefore, every discussion of maximum force implies a clarification of the underlying surface.

Both the force and the power limits state that the flow of momentum or of energy through any *physical surface* – a surface to which an observed can be attached at every one of its points – of any size, for any observer, in any coordinate system, never exceeds the limit value. In particular:

▷ The force limit is only realized *at horizons*. The power limit is only realized *with the help of horizons*.

In all other situations, the observed values are strictly smaller than the maximum values.

The force and power limit values are *invariants*: they are valid for *all* observers and for all interactions. Again, any observer must be a physical system and it must be located Vol. II, page 110 on or near the surface used to define the flow of momentum or energy.

The value of the force limit is the energy of a Schwarzschild black hole divided by its diameter; here the 'diameter' is defined as the circumference divided by π. The power limit is realized when such a black hole is radiated away in the time that light takes to travel along a length corresponding to the diameter.

An object of mass m that has the size of its own Schwarzschild radius $2Gm/c^2$ is called a *black hole*, because according to general relativity, no signals and no light from Ref. 18 inside the Schwarzschild radius can reach the outside world. In this text, black holes are usually non-rotating and usually uncharged; in this case, the terms 'black hole' and 'Schwarzschild black hole' are synonymous.

The value of the maximum force, as well as being the mass–energy of a black hole divided by its diameter, is also the surface gravity of a black hole times its mass. Thus the force limit means that no physical system of a given mass can be concentrated in a region of space-time smaller than a (non-rotating) black hole of that mass. (This is the so-called *hoop conjecture*.) In fact, the mass–energy concentration limit can easily be transformed Challenge 8 e algebraically into the force limit: they are equivalent.

It is easily checked that the maximum force limit is valid for all systems observed in nature, whether they are microscopic, macroscopic or astrophysical. Neither the 'gravitational force' (as long as it is operationally defined) nor the electromagnetic or Challenge 9 e nuclear interactions are ever found to exceed this limit.

But is it possible to *imagine* a system that exceeds the force limit? An extensive discus- Vol. II, page 107 sion shows that this is impossible. For example, the force limit cannot be overcome with Lorentz boosts. We might think that a boost can be chosen in such a way that a 3-force value F in one frame is transformed into any desired value F' in another, boosted frame.

This thought turns out to be wrong. In relativity, 3-force cannot be increased beyond all Vol. II, page 83 bounds using boosts. In all reference frames, the measured 3-force can never exceed the proper force, i.e., the 3-force value measured in the comoving frame.

Also changing to an accelerated frame does not help to overcome the force limit, because for high accelerations a, horizons appear at distance c^2/a, and a mass m has a minimum diameter given by $l \geqslant 4Gm/c^2$.

In fact, the force and power limits cannot be exceeded in any thought experiment, Vol. II, page 107 as long as the sizes of observers or of test masses are taken into account. All apparent exceptions or paradoxes assume the existence of point particles or point-like observers; these, however, are not physical: they do not exist in general relativity.

Fortunately for us, nearby black holes or horizons are rare. Unfortunately, this means that neither the force limit nor the power limit are realized in any physical system at hand, neither at everyday length scales, nor in the microscopic world, nor in astrophysical systems. Even though the force and power limits have never been exceeded, a direct experimental confirmation of the limits will take some time.

The formulation of general relativity as a consequence of a maximum force is not Ref. 19 common; in fact, it seems that it was only discovered 80 years after the theory of general relativity had first been proposed.

Deducing general relativity*

In order to elevate the force or power limit to a principle of nature, we have to show that, just as special relativity follows from the maximum speed, so general relativity follows from the maximum force.

Ref. 20 The maximum force and the maximum power are only realized at horizons. Horizons are regions of space-time where the curvature is so high that it limits the possibility of observation. The name 'horizon' is due to an analogy with the usual horizon of everyday life, which also limits the distance to which we can see. However, in general relativity horizons are *surfaces*, not lines. In fact, we can *define* the concept of horizon in general relativity as a region of maximum force; it is then easy to prove that a horizon is always a two-dimensional surface, and that it is essentially black (except for quantum effects).

The connection between horizons and the maximum force or power allows us to deduce the field equations in a simple way. First, there is always a flow of energy at a horizon. Horizons cannot be planes, since an infinitely extended plane would imply an infinite energy flow. To characterize the finite extension of a given horizon, we use its radius R and its total area A.

The energy flow across a horizon is characterized by an energy E and a proper length L of the energy pulse. When such an energy pulse flows perpendicularly across a horizon, the momentum change $\mathrm{d}p/\mathrm{d}t = F$ is given by

$$F = \frac{E}{L} \,. \tag{11}$$

Since we are at a horizon, we need to insert the maximum possible values. In terms of

* This section can be skipped at first reading.

the horizon area A and radius R, we can rewrite the limit case as

$$\frac{c^4}{4G} = \frac{E}{A} 4\pi R^2 \frac{1}{L} \qquad (12)$$

where we have introduced the maximum force and the maximum possible area $4\pi R^2$ of a horizon of (maximum local) radius R. The ratio E/A is the energy per unit area flowing across the horizon.

Horizons are often characterized by the so-called *surface gravity a* instead of the radius R. In the limit case, two are related by $a = c^2/2R$. This leads to

$$E = \frac{1}{4\pi G} a^2 \, A \, L \, . \qquad (13)$$

Ref. 21 Special relativity shows that at horizons the product aL of proper length and acceleration is limited by the value $c^2/2$. This leads to the central relation for the energy flow at horizons:

$$E = \frac{c^2}{8\pi G} a \, A \, . \qquad (14)$$

This *horizon relation* makes three points. First, the energy flowing across a horizon is limited. Secondly, this energy is proportional to the area of the horizon. Thirdly, the energy flow is proportional to the surface gravity. These three points are fundamental, and characteristic, statements of general relativity. (We also note that due to the limit property of horizons, the energy flow *towards* the horizon just outside it, the energy flow *across* a horizon, and the energy *inside* a horizon are all the same.)

Taking differentials, the horizon relation can be rewritten as

$$\delta E = \frac{c^2}{8\pi G} a \, \delta A \, . \qquad (15)$$

In this form, the relation between energy and area can be applied to general horizons, including those that are irregularly curved or time-dependent.*

Ref. 22 In a well-known paper, Jacobson has given a beautiful proof of a simple connection: if energy flow is proportional to horizon area for all observers and all horizons, and if the proportionality constant is the correct one, then general relativity follows. To see the connection to general relativity, we generalize the horizon relation (15) to general coordinate systems and general directions of energy flow.

* The horizon relation (15) is well known, though with different names for the observables. Since no communication is possible across a horizon, the detailed fate of energy flowing across a horizon is also unknown. Energy whose detailed fate is unknown is often called *heat*, and abbreviated Q. The horizon relation (15) therefore states that the heat flowing through a horizon is proportional to the horizon area. When quantum theory is introduced into the discussion, the area of a horizon can be called 'entropy' S and its surface gravity can be called 'temperature' T; relation (15) can then be rewritten as $\delta Q = T\delta S$. However, this translation of relation (15), which requires the quantum of action, is unnecessary here. We only cite it to show the relation between horizon behaviour and quantum aspects of gravity.

The proof uses tensor notation. We introduce the general surface element $d\Sigma$ and the local boost Killing vector field k that generates the horizon (with suitable norm). We then rewrite the left-hand side of relation (15) as

$$\delta E = \int T_{ab} k^a \, d\Sigma^b \, , \tag{16}$$

where T_{ab} is the energy–momentum tensor. This is valid in arbitrary coordinate systems and for arbitrary energy flow directions. Jacobson's main result is that the right-hand side of the horizon relation (15) can be rewritten, using the (purely geometric) Raychaudhuri equation, as

$$a\, \delta A = c^2 \int R_{ab} k^a \, d\Sigma^b \, , \tag{17}$$

where R_{ab} is the Ricci tensor describing space-time curvature.

Combining these two steps, we find that the energy–area horizon relation (15) can be rewritten as

$$\int T_{ab} k^a \, d\Sigma^b = \frac{c^4}{8\pi G} \int R_{ab} k^a \, d\Sigma^b \, . \tag{18}$$

Jacobson shows that this equation, together with local conservation of energy (i.e., vanishing divergence of the energy–momentum tensor), can only be satisfied if

$$T_{ab} = \frac{c^4}{8\pi G} \left(R_{ab} - \left(\frac{1}{2} R + \Lambda \right) g_{ab} \right) \, , \tag{19}$$

where Λ is a constant of integration whose value is not determined by the problem. These are the full field equations of general relativity, including the cosmological constant Λ. This value of this constant remains undetermined, though.

The field equations are thus shown to be valid at horizons. Now, it is possible, by choosing a suitable coordinate transformation, to position a horizon at any desired space-time event. To achieve this, simply change to the frame of an observer accelerat-
Vol. II, page 98
ing away from that point at the correct distance, as explained in the volume on relativity. Therefore, because a horizon can be positioned anywhere at any time, the field equations must be valid over the whole of space-time.

Since it is possible to have a horizon at every event in space-time, there is the same maximum possible force (or power) at every event in nature. This maximum force (or power) is thus a constant of nature.

In other words, the field equations of general relativity are a direct consequence of the limited energy flow at horizons, which in turn is due to the existence of a maximum force or power. We can thus speak of the maximum force *principle*. Conversely, the field equations imply maximum force and power. Maximum force and general relativity are thus equivalent.

By the way, modern scholars often state that general relativity and gravity follow from the existence of a minimum measurable length. The connection was already stated by
Ref. 23
Sakharov in 1969. This connection is correct, but unnecessarily restrictive. The maximum force, which is implicit in the minimal length, is sufficient to imply gravity. Quantum

theory – or $\hbar$ – is (obviously) not necessary to deduce gravity.

Deducing universal gravitation

Universal gravitation follows from the force limit in the case where both forces and speeds are much smaller than the maximum values. The first condition implies
Challenge 10 e
$\sqrt{4GMa} \ll c^2$, the second $v \ll c$ and $al \ll c^2$. Let us apply this to a specific case. Consider a satellite circling a central mass M at distance R with acceleration a. This system, with length $l = 2R$, has only one characteristic speed. Whenever this speed v is much smaller than c, v^2 must be proportional both to the squared speed calculated by $al = 2aR$ and to the squared speed calculated from $\sqrt{4GMa}$. Taken together, these two
Challenge 11 e
conditions imply that $a = fGM/R^2$, where f is a numerical factor. A quick check, for example using the observed escape velocity values, shows that $f = 1$.

Forces and speeds much smaller than the limit values thus imply that gravity changes with the inverse square of distance. In other words, nature's limit on force implies universal gravitation. Other deductions of universal gravity from limit quantities are given
Page 283
later.

The size of physical systems in general relativity

General relativity, like the other theories of modern physics, implies a limit on the *size* l of systems. There is a limit to the amount of matter that can be concentrated into a small volume:

$$l \geqslant \frac{4Gm}{c^2} . \tag{20}$$

The size limit is only realized for *black holes*, those well-known systems which swallow everything that is thrown into them. The size limit is fully equivalent to the force limit. (Also the hoop conjecture is understood to be true.) All *composite* systems in nature comply with the lower size limit. Whether elementary particles fulfil or even match this
Page 302
limit remains open at this point. This issue will be explored below.

General relativity also implies an 'indeterminacy relation' for the measurement errors
Ref. 24
of size l and energy E of systems:

$$\frac{\Delta E}{\Delta l} \leqslant \frac{c^4}{4G} . \tag{21}$$

Experimental data are available only for composite systems; all known systems
Ref. 25
comply with it. For example, the latest measurements for the Sun give $GM_{\odot}/c^3 = 4.925\,490\,947(1)\,\mu s$; the error in E is thus much smaller than the (scaled) error in its radius, which is known with much smaller precision. The 'indeterminacy relation' (21) is not as well known as that from quantum theory. In fact, tests of it – for example with binary pulsars – may distinguish general relativity from competing theories. We cannot yet say whether this inequality also holds for elementary particles.

A mechanical analogy for the maximum force

The maximum force is central to the theory of general relativity. Indeed, its value (adorned with a factor 2π) appears in the field equations. The importance of the maximum force becomes clearer when we return to our old image of space-time as a deformable mattress. Like any material body, a mattress is described by a material constant that relates the deformation values to the values of applied energy. Similarly, a mattress, like any material, is described by the maximum stress it can bear before it breaks. These two values describe all materials, from crystals to mattresses. In fact, for perfect crystals (without dislocations), these two material constants are the same.

Empty space somehow behaves like a perfect crystal, or a perfect mattress: it has a deformation-energy constant that is equal to the maximum force that can be applied to it. The maximum force describes the elasticity of space-time. The high value of the maximum force tells us that it is difficult to bend space.

Now, materials are not homogeneous: crystals are made up of atoms, and mattresses are made up of foam bubbles. What is the corresponding structure of space-time? This is a central question in the rest of our adventure. One thing is sure: unlike crystals, vacuum has no preferred directions. We now take a first step towards answering the question of the structure of space-time and particles by putting together all the limits found so far.

PLANCK LIMITS FOR ALL PHYSICAL OBSERVABLES

The existence of a maximum force in nature is equivalent to general relativity. As a result, a large part of modern physics can be summarized in four simple and fundamental limit statements on motion:

$$\begin{aligned} &\text{Quantum theory follows from the action limit:} && W \geqslant \hbar \\ &\text{Thermodynamics follows from the entropy limit:} && S \geqslant k \\ &\text{Special relativity follows from the speed limit:} && v \leqslant c \\ &\text{General relativity follows from the force limit:} && F \leqslant \frac{c^4}{4G} \,. \end{aligned} \tag{22}$$

These (corrected) *Planck limits* are valid for all physical systems, whether composite or elementary, and for all observers. Note that the limit quantities of quantum theory, thermodynamics, special and general relativity can also be seen as the right-hand sides of the respective indeterminacy relations. Indeed, the set (4, 7, 9, 21) of indeterminacy relations is fully equivalent to the four limit statements (22).

Challenge 12 e

We note that the different dimensions of the four fundamental limits (22) in nature mean that the four limits are *independent*. For example, quantum effects cannot be used to overcome the force limit; similarly, the power limit cannot be used to overcome the speed limit. There are thus four independent limits on motion in nature.

By combining the four fundamental limits, we can obtain limits on a number of physical observables. The following limits are valid generally, for both composite and elementary systems:

$$\text{time interval:} \qquad t \geqslant \sqrt{\frac{4G\hbar}{c^5}} = 1.1 \cdot 10^{-43}\,\mathrm{s} \tag{23}$$

$$\text{time-distance product:} \quad td \geqslant \frac{4G\hbar}{c^4} = 3.5 \cdot 10^{-78}\,\text{ms} \tag{24}$$

$$\text{acceleration:} \quad a \leqslant \sqrt{\frac{c^7}{4G\hbar}} = 2.8 \cdot 10^{51}\,\text{m/s}^2 \tag{25}$$

$$\text{angular frequency:} \quad \omega \leqslant 2\pi\sqrt{\frac{c^5}{2G\hbar}} = 5.8 \cdot 10^{43}\,\text{/s}\,. \tag{26}$$

Adding the knowledge that space and time can mix, we get

$$\text{distance:} \quad d \geqslant \sqrt{\frac{4G\hbar}{c^3}} = 3.2 \cdot 10^{-35}\,\text{m} \tag{27}$$

$$\text{area:} \quad A \geqslant \frac{4G\hbar}{c^3} = 1.0 \cdot 10^{-69}\,\text{m}^2 \tag{28}$$

$$\text{volume:} \quad V \geqslant \left(\frac{4G\hbar}{c^3}\right)^{3/2} = 3.4 \cdot 10^{-104}\,\text{m}^3 \tag{29}$$

$$\text{curvature:} \quad K \leqslant \frac{c^3}{4G\hbar} = 1.0 \cdot 10^{69}\,\text{/m}^2 \tag{30}$$

$$\text{mass density:} \quad \rho \leqslant \frac{c^5}{16G^2\hbar} = 3.2 \cdot 10^{95}\,\text{kg/m}^3\,. \tag{31}$$

Of course, speed, action, angular momentum, entropy, power and force are also limited, as already stated. The limit values are deduced from the commonly used Planck values simply by substituting $4G$ for G. These limit values are the true *natural units* of nature. In fact, the ideal case would be to redefine the usual Planck values for all observables to these extremal values, by absorbing the numerical factor 4 into the respective definitions. In the following, we call the limit values the *corrected Planck units* or *corrected Planck limits* and assume that the numerical factor 4 has been properly included. In other words:

▷ Every natural unit or (corrected) Planck unit is the limit value of the corresponding physical observable.

Page 59 Most of these limit statements are found scattered throughout the research literature, though the numerical factors often differ. Each limit has attracted a string of publications. Ref. 26 The existence of a smallest measurable distance and time interval of the order of the Planck values is discussed in all approaches to quantum gravity. The maximum curvature Ref. 27 has been studied in quantum gravity; it has important consequences for the 'beginning' of the universe, where it excludes any infinitely large or small observable. The maximum mass density appears regularly in discussions on the energy of the vacuum.

In the following, we often call the collection of Planck limits the *Planck scales*. We will discover shortly that at Planck scales, nature differs in many ways from what we are used to at everyday scales.

> Die Frage über die Gültigkeit der Voraussetzungen der Geometrie im Unendlichkleinen hängt zusammen mit der Frage nach dem innern Grunde der Massverhältnisse des Raumes. Bei dieser Frage, welche wohl noch zur Lehre vom Raume gerechnet werden darf, kommt die obige Bemerkung zur Anwendung, dass bei einer discreten Mannigfaltigkeit das Princip der Massverhältnisse schon in dem Begriffe dieser Mannigfaltigkeit enthalten ist, bei einer stetigen aber anders woher hinzukommen muss. Es muss also entweder das dem Raume zu Grunde liegende Wirkliche eine discrete Mannigfaltigkeit bilden, oder der Grund der Massverhältnisse ausserhalb, in darauf wirkenden bindenden Kräften, gesucht werden.*
>
> Bernhard Riemann, 1854, *Über die Hypothesen, welche der Geometrie zu Grunde liegen.*

Physics, mathematics and simplicity

The four limits of nature of equation (22) – on action, entropy, speed and force – are astonishing. Above all, the four limits are *simple*. For many decades, a silent assumption has guided many physicists: physics requires *difficult* mathematics, and unification requires even more difficult mathematics.

For example, for over thirty years, Albert Einstein searched with his legendary intensity for the unified theory by exploring more and more complex equations. He did so even on his deathbed!** Also most theoretical physicists in the year 2000 held the prejudice that unification requires difficult mathematics. This prejudice is a consequence of over a century of flawed teaching of physics. Flawed teaching is thus one of the reasons that the search for a unified theory was not successful for so long.

The summary of physics with limit statements shows that nature and physics are *simple*. In fact, the essence of the important physical theories is *extremely* simple: special relativity, general relativity, thermodynamics and quantum theory are each based on a simple inequality.

The summary of a large part of physics with inequalities is suggestive. The summary makes us dream that the description of the remaining parts of physics – gauge fields, elementary particles and the complete theory – might be equally simple. Let us continue to explore where the dream of simplicity leads us to.

Limits to space, time and size

> Those are my principles, and if you don't like them ... well, I have others.
>
> Groucho Marx***

* 'The question of the validity of the hypotheses of geometry in the infinitely small is connected to the question of the foundation of the metric relations of space. To this question, which may still be regarded as belonging to the study of space, applies the remark made above; that in a discrete manifold the principles of its metric relations are given in the notion of this manifold, while in a continuous manifold, they must come from outside. Either therefore the reality which underlies space must form a discrete manifold, or the principles of its metric relations must be sought outside it, in binding forces which act upon it.'

Bernhard Riemann is one of the most important mathematicians. 45 years after this statement, Max Planck confirmed that natural units are due to gravitation, and thus to 'binding forces'.

** Interestingly, he also regularly wrote the opposite, as shown on page 87.

*** Groucho Marx (b. 1890 New York City, d. 1977 Los Angeles), well-known comedian.

We have seen that the four fundamental limits of nature (22) result in a minimum distance and a minimum time interval. As the expressions for the limits shows, these minimum intervals arise directly from the *unification* of quantum theory and relativity: they do not appear if the theories are kept separate. In other terms, unification implies that there is a smallest length in nature. This result is important: the formulation of physics as a set of limit statements shows that *the continuum model of space and time is not completely correct.* Continuity and manifolds are only approximations, valid for large actions, low speeds and small forces. Formulating general relativity and quantum theory with limit statements makes this especially clear.

The existence of a force limit in nature implies that no physical system can be smaller than a Schwarzschild black hole of the same mass. In particular, *point particles do not exist*. The density limit makes the same point. In addition, elementary particles are predicted to be larger than the corrected Planck length. So far, this prediction has not been tested by observations, as the scales in question are so small that they are beyond experimental reach. Detecting the sizes of elementary particles – for example, with electric Page 59 dipole measurements – would make it possible to check all limits directly.

Mass and energy limits

Mass plays a special role in all these arguments. The four limits (22) do not make it possible to extract a limit statement on the mass of physical systems. To find one, we have to restrict our aim somewhat.

The Planck limits mentioned so far apply to *all* physical systems, whether composite or elementary. Other limits apply only to elementary systems. In quantum theory, the distance limit is a size limit only for *composite* systems. A particle is *elementary* if its size l is smaller than any measurable dimension. In particular, it must be smaller than the reduced Compton wavelength:

$$\text{for elementary particles:} \quad l \leqslant \frac{\hbar}{mc} \,. \tag{32}$$

Using this limit, we find the well-known mass, energy and momentum limits that are valid *only* for elementary particles:

$$\begin{aligned}
\text{for (real) elementary particles:} \quad m &\leqslant \sqrt{\frac{\hbar c}{4G}} = 1.1 \cdot 10^{-8}\,\mathrm{kg} = 0.60 \cdot 10^{19}\,\mathrm{GeV}/c^2 \\
\text{for (real) elementary particles:} \quad E &\leqslant \sqrt{\frac{\hbar c^5}{4G}} = 9.8 \cdot 10^{8}\,\mathrm{J} = 0.60 \cdot 10^{19}\,\mathrm{GeV} \\
\text{for (real) elementary particles:} \quad p &\leqslant \sqrt{\frac{\hbar c^3}{4G}} = 3.2\,\mathrm{kg\,m/s} = 0.60 \cdot 10^{19}\,\mathrm{GeV}/c \,.
\end{aligned} \tag{33}$$

These elementary-particle limits are the (corrected) *Planck mass*, *Planck energy* and *Planck momentum*. They were discussed in 1968 by Andrei Sakharov, though with different numerical factors. They are regularly cited in elementary particle theory. All known measurements comply with them. Ref. 28

Virtual particles – a new definition

In fact, there are elementary particles that exceed all three limits that we have encountered so far. Nature does have particles which move faster than light, which show actions below the quantum of action, and which experience forces larger than the force limit.

Vol. II, page 72
We know from special relativity that the virtual particles exchanged in collisions move faster than light. We know from quantum theory that the exchange of a virtual particle
Vol. IV, page 193
implies actions below the minimum action. Virtual particles also imply an instantaneous change of momentum; they thus exceed the force limit.

In short, *virtual particles* exceed all the limits that hold for *real* elementary particles.

Curiosities and fun challenges about Planck limits*

The (corrected) Planck limits are statements about properties of nature. There is no way to measure values exceeding these limits, with any kind of experiment. Naturally, such a claim provokes the search for counter-examples and leads to many paradoxes.

* *

The minimum action may come as a surprise at first, because angular momentum and spin have the same unit as action; and nature contains particles with spin 0 or with spin $1/2\,\hbar$. A minimum action indeed implies a minimum angular momentum. However, the angular momentum in question is *total* angular momentum, including the orbital part with respect to the observer. The measured total angular momentum of a particle is never smaller than $\hbar$, even if the spin is smaller.

* *

In terms of mass flows, the power limit implies that flow of water through a tube is limited in throughput. The resulting limit $\mathrm{d}m/\mathrm{d}t \leqslant c^3/4G$ for the change of mass with time seems to be unrecorded in the research literature of the twentieth century.

* *

A further way to deduce the minimum length using the limit statements which structure this adventure is the following. General relativity is based on a maximum force in nature, or alternatively, on a maximum mass change per time, whose value is given by $\mathrm{d}m/\mathrm{d}t =$
Vol. II, page 107
$c^3/4G$. Quantum theory is based on a minimum action W in nature, given by $\hbar$. Since a distance d can be expressed as

$$d^2 = \frac{W}{\mathrm{d}m/\mathrm{d}t}\ , \tag{34}$$

we see directly that a minimum action and a maximum rate of change of mass imply a minimum distance. In other words, quantum theory and general relativity force us to conclude that *in nature there is a minimum distance.* In other words, *at Planck scales the term 'point in space' has no theoretical or experimental basis.*

* Sections called 'Curiosities' can be skipped at first reading.

* *

With the single-particle limits, the entropy limit leads to an upper limit for temperature:

$$T \leqslant \sqrt{\frac{\hbar c^5}{4Gk^2}} = 0.71 \cdot 10^{32}\,\mathrm{K}\,. \qquad (35)$$

This corresponds to the temperature at which the energy per degree of freedom is given by the (corrected) Planck energy $\sqrt{\hbar c^5/4G}$. A more realistic value would have to take account of the number of degrees of freedom of a particle at Planck energy. This would change the numerical factor. However, no system that is even near this temperature value has been studied yet. Only Planck-size horizons are expected to realize the temperature limit, but nobody has managed to explore them experimentally, so far.

* *

How can the maximum force be determined by gravity alone, which is the weakest interaction? It turns out that in situations near the maximum force, the other interactions are usually negligible. This is the reason why gravity must be included in a unified description of nature.

* *

At first sight, it seems that electric charge can be used in such a way that the acceleration of a charged body towards a charged black hole is increased to a value, when multiplied with the mass, that exceeds the force limit. However, the changes in the horizon for
Challenge 13 e charged black holes prevent this.

* *

The gravitational attraction between two masses never yields force values high enough to exceed the force limit. Why? First of all, masses m and M cannot come closer together than the sum of their horizon radii. Using $F = GmM/r^2$ with the distance r given by the (naive) sum of the two black hole radii as $r = 2G(M+m)/c^2$, we get

$$F \leqslant \frac{c^4}{4G}\frac{Mm}{(M+m)^2}\,, \qquad (36)$$

which is never larger than the force limit. Thus even two attracting black holes cannot exceed the force limit – in the inverse-square approximation of universal gravity. In short, the minimum size of masses means that the maximum force cannot be exceeded.

* *

It is well known that gravity bends space. Therefore, if they are to be fully convincing, our calculation for two attracting black holes needs to be repeated taking into account the curvature of space. The simplest way is to study the force generated by a black hole on a test mass hanging from a wire that is lowered towards a black hole horizon. For an
Ref. 29 *unrealistic point mass*, the force would diverge at the horizon. Indeed, for a point mass m lowered towards a black hole of mass M at (conventionally defined radial) distance d,

the force would be

$$F = \frac{GMm}{d^2\sqrt{1 - \frac{2GM}{dc^2}}} . \qquad (37)$$

This diverges at $d = 0$, the location of the horizon. However, even a test mass cannot be smaller than its own gravitational radius. If we want to reach the horizon with a *realistic* test mass, we need to choose a small test mass m: only a small mass can get near the horizon. For vanishingly small masses, however, the resulting force tends to zero. Indeed, letting the distance tend to the smallest possible value by letting $d = 2G(m + M)/c^2 \rightarrow 2GM/c^2$ requires $m \rightarrow 0$, which makes the force $F(m, d)$ vanish. If on the other hand, we remain away from the horizon and look for the maximum force by using a mass as large as can possibly fit into the available distance (the calculation is straightforward), then again the force limit is never exceeded. In other words, for *realistic* test masses, expression (37) is *never* larger than $c^4/4G$. Taking into account the minimal size of test masses, we thus see that the maximum force is never exceeded in gravitational systems.

* *

An absolute power limit implies a limit on the energy that can be transported per unit time through any imaginable physical surface. At first sight, it may seem that the combined power emitted by two radiation sources that each emit 3/4 of the maximum value should give 3/2 times the maximum value. However, the combination forms a black hole, or at least prevents part of the radiation from being emitted by swallowing it between the two sources.

Challenge 14 e

* *

One possible system that actually achieves the Planck power limit is the final stage of black hole evaporation. But even in this case, the power limit is not exceeded.

Challenge 15 e

* *

Ref. 19

The maximum force limit states that the stress-energy tensor, when integrated over any physical surface, does not exceed the limit value. No such integral, over any physical surface, of any tensor component in any coordinate system, can exceed the force limit, provided that it is measured by a realistic observer, in particular, by an observer with a realistic proper size. The maximum force limit thus applies to any component of any force vector, as well as to its magnitude. It applies to gravitational, electromagnetic, and nuclear forces; and it applies to all realistic observers. It is not important whether the forces are real or fictitious; nor whether we are discussing the 3-forces of Galilean physics or the 4-forces of special relativity. Indeed, the force limit applied to the zeroth component of the 4-force is the power limit.

* *

The power limit is of interest if applied to the universe as a whole. Indeed, it can be used to partly explain Olbers' paradox: the sky is dark at night because the combined luminosity of all light sources in the universe cannot be brighter than the maximum value.

* *

Page 36 The force limit and its solid state analogy might be seen to suggest that the appearance of matter might be nature's way of preventing space from ripping apart. Does this analogy make sense? Challenge 16 s

* *

Ref. 23 In fact, the connection between minimum length and gravity is not new. Already in 1967, Andrei Sakharov pointed out that a minimum length implies gravity. He showed that regularizing quantum field theory on curved space with a cut-off at small distances will induce counter-terms that include to lowest order the cosmological constant and then the Einstein–Hilbert action of general relativity.

* *

We said above that a surface is physical if an observer can be attached to each of its points. The existence of a smallest length – and a corresponding shortest time interval – implies

▷ No surface is *physical* if any part of it requires a localization in space-time to scales below the minimum length.

For example, a physical surface must not cross any horizon. Only by insisting on physical surfaces can we eliminate unphysical examples that contravene the force and power limits. For example, this condition was overlooked in Bousso's early discussion of Bekenstein's entropy bound – though not in his more recent ones. Ref. 30

* *

The equation $E = c^2 m$ implies that energy and mass are equivalent. What do the equations $l = (4G/c^2)m = (4G/c^4)E$ for length and $W = \hbar\varphi$ for action imply? Challenge 17 e

* *

Our discussion of limits can be extended to include electromagnetism. Using the (low-energy) electromagnetic coupling constant α, the fine structure constant, we get the following limits for physical systems interacting electromagnetically:

$$\text{electric charge:} \quad q \geqslant \sqrt{4\pi\varepsilon_0 \alpha c \hbar} = e = 0.16\,\mathrm{aC} \tag{38}$$

$$\text{electric field:} \quad E \leqslant \sqrt{\frac{c^7}{64\pi\varepsilon_0 \alpha \hbar G^2}} = \frac{c^4}{4Ge} = 1.9 \cdot 10^{62}\,\mathrm{V/m} \tag{39}$$

$$\text{magnetic field:} \quad B \leqslant \sqrt{\frac{c^5}{64\pi\varepsilon_0 \alpha \hbar G^2}} = \frac{c^3}{4Ge} = 6.3 \cdot 10^{53}\,\mathrm{T} \tag{40}$$

$$\text{voltage:} \quad U \leqslant \sqrt{\frac{c^4}{16\pi\varepsilon_0 \alpha G}} = \frac{1}{e}\sqrt{\frac{\hbar c^5}{4G}} = 6.1 \cdot 10^{27}\,\mathrm{V} \tag{41}$$

$$\text{inductance:} \quad L \geqslant \frac{1}{4\pi\varepsilon_0 \alpha}\sqrt{\frac{4G\hbar}{c^7}} = \frac{1}{e^2}\sqrt{\frac{4G\hbar^3}{c^5}} = 4.4 \cdot 10^{-40}\,\mathrm{H}\,. \tag{42}$$

With the additional assumption that in nature at most one particle can occupy one Planck volume, we get

$$\text{charge density:} \qquad \rho_e \leqslant \sqrt{\frac{\pi\varepsilon_0\alpha}{16G^3}}\,\frac{c^5}{\hbar} = e\sqrt{\frac{c^9}{64G^3\hbar^3}} = 4.7\cdot10^{84}\ \mathrm{C/m^3} \qquad (43)$$

$$\text{capacitance:} \qquad C \geqslant 4\pi\varepsilon_0\alpha\sqrt{\frac{4G\hbar}{c^3}} = e^2\sqrt{\frac{4G}{c^5\hbar}} = 2.6\cdot10^{-47}\ \mathrm{F}\,. \qquad (44)$$

For the case of a single conduction channel, we get

$$\text{electric resistance:} \qquad R \geqslant \frac{1}{4\pi\varepsilon_0\alpha c} = \frac{\hbar}{e^2} = 4.1\ \mathrm{k\Omega} \qquad (45)$$

$$\text{electric conductivity:} \qquad G \leqslant 4\pi\varepsilon_0\alpha c = \frac{e^2}{\hbar} = 0.24\ \mathrm{mS} \qquad (46)$$

$$\text{electric current:} \qquad I \leqslant \sqrt{\frac{\pi\varepsilon_0\alpha c^6}{G}} = e\sqrt{\frac{c^5}{4\hbar G}} = 1.5\cdot10^{24}\ \mathrm{A}\,. \qquad (47)$$

The magnetic field limit is significant in the study of extreme stars and black holes. The maximum electric field plays a role in the theory of gamma-ray bursters. For current, conductivity and resistance in single channels, the limits and their effects were studied extensively in the 1980s and 1990s.

Ref. 31

Ref. 32

The observation of quarks and of collective excitations in semiconductors with charge $e/3$ does not necessarily invalidate the charge limit for physical systems. In neither case is there is a physical system – defined as localized mass–energy interacting incoherently with the environment – with charge $e/3$.

* *

The general relation that to every limit value in nature there is a corresponding indeterminacy relation is valid also for electricity. Indeed, there is an indeterminacy relation for capacitors, of the form

$$\Delta C\ \Delta U \geqslant e\,, \qquad (48)$$

where e is the positron charge, C capacity and U potential difference. There is also an indeterminacy relation between electric current I and time t

$$\Delta I\ \Delta t \geqslant e\,. \qquad (49)$$

Ref. 33 Both these relations may be found in the research literature.

Cosmological limits for all physical observables

In our quest to understand motion, we have focused our attention on the four fundamental limitations to which motion is subject. Special relativity posits a limit to speed, namely the speed of light c. General relativity limits force and power respectively by $c^4/4G$ and $c^5/4G$, and quantum theory introduces a smallest value $\hbar$ for action. Nature imposes the lower limit k on entropy. If we include the limit e on electric charge changes, these limits induce extremal values for *all* physical observables, given by the corresponding (corrected) Planck values.

Vol. II, page 107

A question arises: does nature also impose limits on physical observables at the opposite end of the measurement scale? For example, there is a highest force and a highest power in nature. Is there also a lowest force and a lowest power? Is there also a lowest speed? We will show that there are indeed such limits, for all observables. We give the general method to generate such bounds, and explore several examples. This exploration will take us on an interesting survey of modern physics. We start by deducing system-dependent limits and then go on to cosmological limits.

Size and energy dependence

While looking for additional limits in nature, we note a fundamental fact. Any upper limit for angular momentum, and any lower limit for power, must be *system-dependent*. Such limits will not be absolute, but will depend on properties of the system. Now, a physical system is a part of nature characterized by a boundary and its content.* Thus the simplest properties shared by all systems are their size (characterized in the following by the diameter) L and their energy E. With these characteristics we can deduce system-dependent limits for every physical observable. The general method is straightforward: we take the known inequalities for speed, action, power, charge and entropy, and then extract a limit for any observable, by inserting the length and energy as required. We then have to select the strictest of the limits we find.

Angular momentum and action

The ratio of angular momentum D to energy E times length L has the dimensions of inverse speed. Since rotation speeds are limited by the speed of light, we get

Challenge 18 e

$$D_{\text{system}} \leqslant \frac{1}{c} LE \,. \tag{50}$$

Indeed, in nature there do not seem to be any exceptions to this limit on angular momentum. In no known system, from atoms to molecules, from ice skaters to galaxies, does the angular momentum exceed this value. Even the most violently rotating objects, the so-called extremal black holes, are limited in angular momentum by $D \leqslant LE/c$. (Actually, this limit is correct for black holes only if the energy is taken as the irreducible mass times c^2; if the usual mass is used, the limit is too large by a factor of 4.) The limit

Ref. 34

* Quantum theory refines this definition: a physical system is a part of nature that in addition interacts *incoherently* with its environment. In the following discussion we will assume that this condition is satisfied.

Vol. IV, page 151

deduced from general relativity, given by $D \leqslant L^2c^3/4G$, is not stricter than the one just given. By the way, no system-dependent lower limit for angular momentum can be deduced.

The maximum value for angular momentum is also interesting when it is seen as an action limit. Action is the time integral of the difference between kinetic and potential energy. Though nature always seeks to minimize the action W, systems, of size L, that *maximize* action are also interesting. You might check for yourself that the action limit

$$W \leqslant LE/c \tag{51}$$

Challenge 19 e is not exceeded in any physical process.

Speed

Speed times mass times length is an action. Since action values in nature are limited from below by $\hbar$, we get a limit for the speed of a system:

$$v_{\text{system}} \geqslant \hbar c^2 \frac{1}{LE} . \tag{52}$$

This is not a new result; it is just a form of the indeterminacy relation of quantum theory. It gives a minimum speed for any system of energy E and diameter L. Even the extremely slow radius change of a black hole by evaporation just realizes this minimal speed. Challenge 20 e

Continuing with the same method, we also find that the limit deduced from general relativity, $v \leqslant (c^2/4G)(L/E)$, gives no new information. Therefore, no *system-dependent* upper speed limit exists – just the global limit c. Challenge 21 e

Incidentally, the limits are not unique. Other limits can be found in a systematic way. Upper limits can be multiplied, for example, by factors of $(L/E)(c^4/4G)$ or $(LE)(2/\hbar c)$, yielding less strict upper limits. A similar rule can be given for lower limits. Challenge 22 s

Force, power and luminosity

We have seen that force and power are central to general relativity. The force exerted by a system is the flow of momentum out of the system; emitted power is the flow of energy out of the system. Thanks to the connection $W = FLT$ between action W, force F, distance L and time T, we can deduce

$$F_{\text{system}} \geqslant \frac{\hbar}{2c} \frac{1}{T^2} . \tag{53}$$

Experiments do not reach this limit. The smallest forces measured in nature are those in atomic force microscopes, where values as small as 1 aN are observed. But even these values are above the lower force limit.

The power P emitted by a system of size L and mass M is limited by

$$c^3 \frac{M}{L} \geqslant P_{\text{system}} \geqslant 2\hbar G \frac{M}{L^3} . \tag{54}$$

The limit on the left is the upper limit for any engine or lamp, as deduced from relativity; not even the universe exceeds it. The limit on the right is the minimum power emitted by any system through quantum gravity effects. Indeed, no physical system is completely tight. Even black holes, the systems with the best ability to keep components inside their enclosure, radiate. The power radiated by black holes should just meet this limit, provided the length L is taken to be the circumference of the black hole. Thus the claim of the quantum gravity limit is that the power emitted by a black hole is the smallest power that is emitted by any composite system of the same surface gravity. (However, the numerical factors in the black hole power appearing in the research literature are not yet consistent.)

The strange charm of the entropy bound

Ref. 35 In 1973, Bekenstein discovered a famous limit that connects the entropy S of a physical system with its size and mass. No system has a larger entropy than one bounded by a horizon. The larger the horizon surface, the larger the entropy. We write

$$\frac{S}{S_{\text{c.Planck}}} \leqslant \frac{A}{A_{\text{c.Planck}}} \tag{55}$$

which gives

$$S \leqslant \frac{kc^3}{4G\hbar} A , \tag{56}$$

where A is the surface of the system. Equality is realized only for black holes. The old question of the origin of the factor 4 in the entropy of black holes is thus answered here: it is due to the factor 4 in the force or power bound in nature. Time will tell whether this explanation will be generally accepted.

We can also derive a more general relation by using a mysterious assumption, which we will discuss afterwards. We assume that the limits for vacuum are opposite to those for matter. We can then write $c^2/4G \leqslant M/L$ for the vacuum. Using

$$\frac{S}{S_{\text{c.Planck}}} \leqslant \frac{M}{M_{\text{c.Planck}}} \frac{A}{A_{\text{c.Planck}}} \frac{L_{\text{c.Planck}}}{L} \tag{57}$$

we get

$$S \leqslant \frac{\pi kc}{\hbar} ML = \frac{2\pi kc}{\hbar} MR . \tag{58}$$

This is called *Bekenstein's entropy bound*. It states that the entropy of any physical system is *finite* and limited by its mass M and size L. No exception has ever been found or
Ref. 30 constructed, despite many attempts. Again, the limit value itself is only realized for black holes.

We need to explain the strange assumption used above. We are investigating the entropy of a horizon. Horizons are not matter, but limits to empty space. The entropy of horizons is due to the large number of virtual particles found at them. In order to deduce the maximum entropy of expression (57) we therefore have to use the properties of the vacuum. In other words, *either* we use a mass-to-length ratio for vacuum *above* the Planck limit, *or* we use the Planck entropy as the *maximum* value for vacuum.

Ref. 36 Other, equivalent limits for entropy can be found if other variables are introduced. For example, since the ratio of the shear viscosity η to the volume density of entropy (times k) has the dimensions of action, we can directly write

$$S \leqslant \frac{k}{\hbar}\eta V . \qquad (59)$$

Again, equality is only attained in the case of black holes. In time, no doubt, the list of similar bounds will grow longer.

Is there also a smallest, system-dependent entropy? So far, there does not seem to be a system-dependent minimum value for entropy: the present approach gives no expression that is larger than k. Challenge 23 e

The establishment of the entropy limit is an important step towards making our description of motion consistent. If space-time can move, as general relativity maintains, it also has an entropy. How could entropy be finite if space-time were continuous? Clearly, because of the existence of a minimum distance and minimum time in nature, space-time cannot be continuous, but must have a finite number of degrees of freedom, and thus a finite entropy.

Curiosities and fun challenges about system-dependent limits to observables

Like the Planck values, also the system-dependent limit values for all physical observables yield a plethora of interesting questions. We study a few examples.

∗ ∗

The content of a system is characterized not only by its mass and charge, but also by its strangeness, isospin, colour charge, charge and parity. Can you deduce the limits for these quantities? Challenge 24 r

∗ ∗

In our discussion of black hole limits, we silently assumed that they interact, like any thermal system, in an incoherent way with the environment. Which of the results of this section change when this condition is dropped, and how? Which limits can be overcome? Challenge 25 s

∗ ∗

Challenge 26 e Can you find a general method to deduce all limits of observables?

∗ ∗

Bekenstein's entropy bound leads to some interesting speculations. Let us speculate that the universe itself, being surrounded by a horizon, meets the Bekenstein bound. The entropy bound gives a bound to all degrees of freedom inside a system: it tells us that the

Challenge 27 e number $N_{\text{d.o.f.}}$ of degrees of freedom in the universe is roughly

$$N_{\text{d.o.f.}} \approx 10^{132} . \tag{60}$$

Compare this with the number $N_{\text{Pl. vol.}}$ of Planck volumes in the universe

$$N_{\text{Pl. vol.}} \approx 10^{183} \tag{61}$$

and with the number $N_{\text{part.}}$ of particles in the universe

$$N_{\text{part.}} \approx 10^{91} . \tag{62}$$

We see that particles are only a tiny fraction of what moves around. Most motion must be movement of space-time. At the same time, space-time moves far less than might be naively expected. To find out how all this happens is the challenge of the unified description of motion.

* *

A lower limit for the temperature of a thermal system can be found using the following idea: the number of degrees of freedom of a system is limited by its surface, or more precisely, by the ratio between the surface and the Planck surface. We get the limit

$$T \geqslant \frac{4G\hbar}{\pi k c} \frac{M}{L^2} . \tag{63}$$

This is the smallest temperature that a system of mass M and size L can have. Alternatively, using the method given above, we can use the limit on the thermal energy $kT/2 \geqslant \hbar c/2\pi L$ (the thermal wavelength must be smaller than the size of the system) together with the limit on mass $c^2/4G \geqslant M/L$, and deduce the same result.

We have met the temperature limit already: when the system is a black hole, the limit yields the temperature of the emitted radiation. In other words, the temperature of black holes is the lower limit for all physical systems for which a temperature can be defined, provided they share the *same boundary gravity*. The latter condition makes sense: boundary gravity is accessible from the outside and describes the full physical system, since it depends on both its boundary and its content.

So far, no exception to the claim on the minimum system temperature is known. All systems from everyday life comply with it, as do all stars. Also the coldest known systems in the universe, namely Bose–Einstein condensates and other cold gases produced Challenge 28 s in laboratories, are much hotter than the limit, and thus much hotter than black holes Vol. II, page 62 of the same surface gravity. (We saw earlier that a consistent Lorentz transformation for temperature is not possible; so the minimum temperature limit is only valid for an observer at the same gravitational potential as the system under consideration and stationary relative to it.)

By the way, there seems to be no consistent way to define an upper limit for a size-Challenge 29 ny dependent temperature. Limits for other thermodynamic quantities can be found, but we do not discuss them here.

* *

When electromagnetism plays a role in a system, the system also needs to be characterized by a charge Q. Our method then gives the following limit for the electric field E:

$$E \geqslant 4Ge \, \frac{M^2}{Q^2 L^2} \; . \tag{64}$$

We write the field limit in terms of the elementary charge e, though it might be more appropriate to write it using the fine structure constant via $e = \sqrt{4\pi\varepsilon_0 \alpha \hbar c}$. In observations, the electric field limit has never been exceeded. For the magnetic field we get

$$B \geqslant \frac{4Ge}{c} \, \frac{M^2}{Q^2 L^2} \; . \tag{65}$$

Again, this limit is satisfied by all known systems in nature.

Similar limits can be found for the other electromagnetic observables. In fact, several of the limits given earlier are modified when electric charge is included. Does the size limit change when electric charge is taken into account? In fact, an entire research field is dedicated to deducing and testing the most general limits valid in nature.

Challenge 30 s

* *

Many cosmological limits have not been discussed here nor anywhere else. The following could all be worth a publication: What is the limit for momentum? Energy? Pressure? Acceleration? Mass change? Lifetime?

Simplified cosmology in one statement

We now continue our exploration of limits by focussing on the largest systems possible in nature. In order to do that, we have a simplified look at cosmology.

The dark sky at night tells us there is a cosmological horizon. This implies

▷ There is a maximum distance value in nature, given by the radius of the cosmological horizon R_c.

For all systems and all observers, sizes, distances and lengths l are observed to be limited by

$$l \lesssim R_C \approx 4.3 \times 10^{26}\,\mathrm{m} = 4.4 \times 10^{10}\,\mathrm{al} \approx 3.3 c t_0 \; . \tag{66}$$

The cosmological horizon limits system sizes. The observed radius is about 3.3 times the age of the universe t_0, times c.

The cosmological horizon also limits the age of systems.

The cosmological limits to observables

From the system-dependent limits for speed, action, force and entropy we can deduce system-dependent limits for all other physical observables. In addition, we note that the system-dependent limits can (usually) be applied to the universe as a whole; we only need to insert the size and energy content of the universe. Usually, we can do this through a limit process, even though the universe itself is not a physical system. In this way, we get an absolute limit for every physical observable that contains the cosmological radius R_C. These limits are on the opposite end of the Planck limit for each observable. We call these limits the *cosmological limits*.

The simplest cosmological limit is the upper limit to length in the universe. Since the cosmological length limit also implies a maximum possible Compton wavelength, we get a minimum particle mass and energy. We also get an cosmological lower limit on luminosity.

For single particles, we find an absolute lower speed limit, the *cosmological speed limit*, given by

$$v_{\text{particle}} \geqslant \frac{L_{\text{c. Planck}}}{R_C} c \approx 7 \cdot 10^{-53}\ \text{m/s}\ . \tag{67}$$

This speed value has never been reached or approached by any observation.

Many cosmological limits are related to black hole limits. For example, the observed average mass density of the universe is not far from the corresponding black hole limit. As another example, the black hole lifetime limit might be imagined to provide an upper limit for the full lifetime of the universe. However, the age of the universe is far from that

Challenge 31 e

limit by a large factor. In fact, since the universe's size and age are increasing, the lifetime limit is pushed further into the future with every second that passes. The universe could be said to evolve so as to escape its own decay...

Minimum force

The negative energy volume density $-\Lambda c^4/4\pi G$ introduced by the positive cosmological constant Λ corresponds to a negative pressure (both quantities have the same dimensions). When multiplied by the minimum area it yields a force value

$$F = \frac{\Lambda \hbar c}{2\pi} = 4.8 \cdot 10^{-79}\ \text{N}\ . \tag{68}$$

Apart from the numerical factor, this is the *cosmological force limit*, the smallest possible force in nature. This is also the gravitational force between two corrected Planck masses located at the cosmological distance $\sqrt{\pi/4\Lambda} \approx R_C$.

As a note, this leads to the fascinating conjecture that the full theory of general relativity, including the cosmological constant, might be defined by the combination of a maximum and a minimum force in nature.

Summary on cosmological limits

The above short exploration of cosmological limits could be extended. Overall, it appears

▷ Nature provides two limits for each observable: a Planck limit and a cosmological limit.

Every observable has a lower and an upper limit. You may want to summarize them into
Challenge 32 s
a table.

All these limits only appear when quantum theory and gravity are brought together. But the existence of these limits, and in particular the existence of limits to measurement precision, forces us to abandon some cherished assumptions.

Limits to measurement precision

We now know that in nature, every physical measurement has a lower and an upper bound. One of the bounds is cosmological, the other is given by the (corrected) Planck unit. As a consequence, for every observable, the smallest relative measurement error that is possible in nature is the ratio between the smaller and the larger limit. In particular, we have to conclude that *all measurements are limited in precision.*

No real numbers

Because of the fundamental limits to measurement precision, *the measured values of physical observables do not require the full set of real numbers.* In fact, limited precision implies that observables cannot be described by the real numbers! This staggering result appears whenever quantum theory and gravity are brought together. But there is more.

Vacuum and mass: two sides of the same coin

There is a limit to the precision of length measurements in nature. This limit is valid both for length measurements of empty space and for length measurements of matter (or radiation). Now let us recall what we do when we measure the length of a table with a ruler. To find the ends of the table, we must be able to distinguish the table from the surrounding air. In more precise terms, we must be able to distinguish matter from vacuum.

Whenever we want high measurement precision, we need to approach Planck scales. But at Planck scales, the measurement values and the measurement errors are of the same size. In short, at Planck scales, the intrinsic measurement limitations of nature imply that we cannot say whether we are measuring vacuum or matter. We will check this
Page 59
conclusion in detail later on.

In fact, we can pick any other observable that distinguishes vacuum from matter – for example, colour, mass, size, charge, speed or angular momentum – and we have the same problem: at Planck scales, the limits to observables lead to limits to measurement precision, and therefore, at Planck scales it is impossible to distinguish between matter and vacuum. At Planck scales, we cannot tell whether a box is full or empty.

To state the conclusion in the sharpest possible terms: *vacuum and matter do not differ at Planck scales*. This counter-intuitive result is one of the charms of the search for a complete, unified theory. It has inspired many researchers in the field and some have
Ref. 37
written best-sellers about it.

The conclusion of indistinguishability of matter and vacuum also arises at cosmological scale: it is intrinsic to the properties of a horizon.

No points

Limited measurement precision also implies that at the Planck energy it is impossible to speak about points, instants, events or dimensionality. Similarly, at the Planck length it is impossible to distinguish between positive and negative time values: so particles and antiparticles are not clearly distinguished at Planck scales. All these conclusions are so far-reaching that we must check them in more detail. We will do this shortly. Page 82

Measurement precision and the existence of sets

In physics, it is generally assumed that nature is a *set* of components or parts. These components, called *elements* by mathematicians, are assumed to be *separable* from each other. This tacit assumption is introduced in three main situations: it is assumed that matter consists of separable particles, that space-time consists of separable events or points, and that the set of states consists of separable initial conditions. Until the year 2000, physics has built the whole of its description of nature on the concept of a set.

The existence of a fundamental limit to measurement precision implies that nature is *not* a set of such separable elements. Precision limits imply that physical entities can be distinguished only *approximately*. The approximate distinction is only possible at energies much lower than the Planck energy $\sqrt{\hbar c^5/4G}$. As humans, we do live at such small energies, and we can safely make the approximation. Indeed, the approximation is excellent in practice; we do not notice any error. But at Planck energy, distinction and separation is impossible in principle. In particular, at the cosmic horizon, at the big bang, and at Planck scales, any precise distinction between two events, two points or two particles becomes impossible. Page 59

Another way to reach this result is the following. Separation of two entities requires *different measurement results* – for example, different positions, different masses or different velocities. Whatever observable is chosen, at the Planck energy the distinction becomes impossible because of the large measurements errors. Only at everyday energies is a distinction possible. In fact, even at everyday energies, any distinction between two physical systems – for example, between a toothpick and a mountain – is possible only *approximately*. At Planck scales, a boundary can never be drawn. Page 59

A third argument is the following. In order to *count* any entities in nature – a set of particles, a discrete set of points, or any other discrete set of physical observables – the entities have to be separable. But the inevitable measurement errors contradict separability. Thus at the Planck energy it is impossible to count physical objects with precision: Page 59

▷ Nature has no parts.

In summary, at Planck scales, perfect separation is impossible in principle. We cannot distinguish observations. *At Planck scales it is impossible to split nature into separate parts or entities.* In nature, elements of sets cannot be defined. Neither discrete nor continuous sets can be constructed: Page 59

▷ Nature does not contain sets or elements.

Since sets and elements are only approximations, the concept of a 'set', which assumes separable elements, is *too specialized* to describe nature. Nature cannot be described at Planck scales – i.e., with full precision – if any of the concepts used for its description presupposes sets. However, all concepts used in the past 25 centuries to describe nature – particles, space, time, observables, phase space, wave functions, Hilbert space, Fock space, Riemannian space, particle space, loop space or moduli space – are based on sets. They must all be abandoned at Planck energy.

▷ No correct mathematical model of nature can be based on sets.

This is a central requirement for the complete theory. In other terms, nature has no parts: nature is one.

None of the approaches to unification pursued in the twentieth century has abandoned sets. The requirement about the unified description tells us too search for *other* approaches. But the requirement about the complete description does more than that. Indeed, the requirement to abandon sets will be an efficient *guide* in our search for the unification of relativity and quantum theory. The requirement will even settle Hilbert's sixth problem.

Page 110

SUMMARY ON LIMITS IN NATURE

If we exclude gauge interactions, we can summarize the rest of physics five limit statements:

▷ The speed limit is equivalent to special relativity.
▷ The force limit is equivalent to general relativity.
▷ The action limit is equivalent to quantum theory.
▷ The entropy limit is equivalent to thermodynamics.
▷ The distance limit is equivalent to cosmology.

All these limits are observer-invariant. The invariance of the limits suggests interesting thought experiments, none of which leads to their violation.

The invariant limits imply that in nature every physical observable is bound on one end by the corresponding (corrected) Planck unit and on the other end by a cosmological limit. Every observable in nature has an upper and lower limit value.

The existence of lower and upper limit values to all observables implies that measurement precision is limited. As a consequence, matter and vacuum are indistinguishable, the description of space-time as a continuous manifold of points is not correct, and nature can be described by sets and parts only approximately. At Planck scales, nature does not contain sets or elements.

Nature's limits imply that Planck units are the key to the unified theory. Since the most precise physical theories known, quantum theory and general relativity, can be reduced to limit statements, there is a good chance that the complete, unified theory of physics will allow an equally simple description. Nature's limits thus suggest that the mathematics of the complete, unified theory might be *simple*.

At this point of our adventure, many questions are still open. Answering any of the Page 19 open issues of the millennium list still seems out of reach. But this impression is too pessimistic. Our discussion implies that we only need to find a description of nature that is simple and *without sets*. And a natural way to avoid the use of sets is a description of empty space, radiation and matter as being made of *common* constituents. But before we explore this option, we check the conclusions of this chapter in another way. In particular, as a help to more conservative physicists, we check all conclusions we found so far *without* making use of the maximum force principle.

Chapter 3

GENERAL RELATIVITY VERSUS QUANTUM THEORY

> “Man muß die Denkgewohnheiten durch Denknotwendigkeiten ersetzen.** ”
>
> Albert Einstein

The two accurate descriptions of motion available in the year 2000, namely that of general relativity and that of the standard model, are both useful and thoroughly beautiful. This millennium description of motion is *useful* because its consequences are confirmed by all experiments, to the full measurement precision. We are able to describe and understand all examples of motion that have ever been encountered. We can use this understanding to save lives, provide food and enjoy life. We have thus reached a considerable height in our mountain ascent. Our quest for the full description of motion is not far from completion.

The results of twentieth century physics are also *beautiful*. By this, physicists just mean that they can be phrased in *simple* terms. This is a poor definition of beauty, but physicists are rarely experts on beauty. Nevertheless, if a physicist has some other concept of beauty in physics, avoid him, because in that case he is really talking nonsense.

The simplicity of twentieth-century physics is well-known: all motion observed in nature minimizes action. Since in physics, action is a measure of change, we can say that all motion observed in nature *minimizes change*. In particular, every example of motion due to general relativity or to the standard model of particle physics minimizes action:
Vol. I, page 248
both theories can be described concisely with the help of a Lagrangian.

On the other hand, some important aspects of any type of motion, the masses of the involved elementary particles and the strength of their coupling, are unexplained by general relativity and by the standard model of particle physics. The same applies to the origin of all the particles in the universe, their initial conditions, and the dimensionality of space-time. Obviously, the millennium description of physics is not yet complete.

The remaining part of our adventure will be the most demanding. In the ascent of any high mountain, the head gets dizzy because of the lack of oxygen. The finite amount of energy at our disposal requires that we leave behind all unnecessary baggage and everything that slows us down. In order to determine what is unnecessary, we need to focus on what we want to achieve. Our aim is the precise description of motion. But even though general relativity and quantum theory are extremely precise, useful and simple, we do carry a burden: the two theories and their concepts *contradict* each other.

** ‘One needs to replace habits of thought by necessities of thought.’

The contradictions

In classical physics and in general relativity, the vacuum, or empty space, is a region with no mass, no energy and no momentum. If particles or gravitational fields are present, the energy density is not zero, space is curved and there is no complete vacuum.

In everyday life, vacuum has an energy density that cannot be distinguished from zero. However, general relativity proposes a way to check this with high precision: we
Ref. 38 measure the average curvature of the universe. Nowadays, cosmological measurements performed with dedicated satellites reveal an average energy density E/V of the inter-
Vol. II, page 243 galactic 'vacuum' with the value of

$$\frac{E}{V} \approx 0.5\,\mathrm{nJ/m^3}\ . \qquad (69)$$

In short, cosmological data show that the energy density of intergalactic space is not exactly zero; nevertheless, the measured value is extremely small and can be neglected in all laboratory experiments.

Ref. 39 On the other hand, quantum field theory tells a different story on vacuum energy density. A vacuum is a region with zero-point fluctuations. The energy content of a vacuum is the sum of the zero-point energies of all the fields it contains. Indeed, the Casimir
Vol. V, page 123 effect 'proves' the reality of these zero-point energies. Following quantum field theory, the most precise theory known, their energy density is given, within one order of mag-
Vol. V, page 129 nitude, by

$$\frac{E}{V} \approx \frac{4\pi h}{c^3}\int_0^{\nu_{\max}} \nu^3 \mathrm{d}\nu = \frac{\pi h}{c^3}\nu_{\max}^4\ . \qquad (70)$$

The approximation is valid for the case in which the cut-off frequency $\nu_{\max}$ is much larger than the rest mass m of the particles corresponding to the field under consideration. The
Page 40 limit considerations given above imply that the cut-off energy has to be of the order of the Planck energy $\sqrt{\hbar c^5/4G}$, about $0.6 \cdot 10^{19}$ GeV= 1.0 GJ. That would give a vacuum energy density of

$$\frac{E}{V} \approx 10^{111}\,\mathrm{J/m^3}\ , \qquad (71)$$

which is about 10^{120} times higher than the experimental measurement. In other words, something is slightly wrong in the calculation due to quantum field theory.*

General relativity and quantum theory contradict each other in other ways. Gravity
Ref. 41 is curved space-time. Extensive research has shown that quantum field theory, which describes electrodynamics and nuclear forces, fails for situations with strongly curved space-time. In these cases the concept of 'particle' is not precisely defined. Quantum field theory cannot be extended to include gravity consistently, and thus to include general relativity. Without the concept of the particle as a discrete entity, we also lose the ability to perform perturbation calculations – and these are the only calculations possible

* It is worthwhile to stress that the 'slight' mistake lies in the domain of quantum field theory. There is no mistake and no mystery, despite the many claims to the contrary found in newspapers and in bad research
Ref. 40 articles, in general relativity. This well-known point is made especially clear by Bianchi and Rovelli.

in quantum field theory. In short, quantum theory only works because it assumes that gravity does not exist. Indeed, the gravitational constant G does not appear in quantum field theory.

On the other hand, general relativity neglects the commutation rules between physical quantities discovered in experiments on a microscopic scale. General relativity assumes that the classical notions of position and momentum of material objects are meaningful. It thus ignores Planck's constant $\hbar$, and only works by neglecting quantum effects.

The concept of *measurement* also differs. In general relativity, as in classical physics, Vol. V, page 295 it is assumed that arbitrary precision of measurement is possible – for example, by using finer and finer ruler marks. In quantum mechanics, on the other hand, the precision of measurement is limited. The indeterminacy relation yields limits that follow from the Vol. V, page 45 mass M of the measurement apparatus.

The contradictions also concern the concept of *time*. According to relativity and classical physics, time is what is read from clocks. But quantum theory says that precise clocks do not exist, especially if gravitation is taken into account. What does 'waiting 10 minutes' mean, if the clock goes into a quantum-mechanical superposition as a result of Page 65 its coupling to space-time geometry? It means nothing.

Similarly, general relativity implies that space and time cannot be distinguished, whereas quantum theory implies that matter does make a distinction between them. A related difference is the following. Quantum theory is a theory of – admittedly weird – Vol. II, page 284 local observables. In general relativity, there are no local observables, as Einstein's hole argument shows.

The contradiction between the two theories is shown most clearly by the failure of general relativity to describe the pair creation of particles with spin 1/2, a typical and Ref. 42, Ref. 43 essential quantum process. John Wheeler* and others have argued that, in such a case, the Ref. 44, Ref. 45 topology of space necessarily has to *change*; in general relativity, however, the topology of space is fixed. Equivalently, quantum theory says that matter is made of *fermions*, but fermions cannot be incorporated into general relativity.**

Ref. 46 Another striking contradiction was pointed out by Jürgen Ehlers. Quantum theory is built on point particles, and point particles move on *time-like* world lines. But following general relativity, point particles have a singularity inside their black hole horizon; and singularities always move on *space-like* world lines. The two theories thus contradict each other at smallest distances.

No description of nature that contains contradictions can lead to a unified or to a completely correct description. To eliminate the contradictions, we need to understand their origin.

The origin of the contradictions

Ref. 47 *All* contradictions between general relativity and quantum mechanics have the same origin. In 20th-century physics, motion is described in terms of objects, made up of

* John Archibald Wheeler (b. 1911, Jacksonville, d. 2008, Hightstown), was a physicist and influential teacher who worked on general relativity.

** As we will see below, the strand conjecture provides a way to incorporate fermions into an extremely accurate approximation of general relativity, without requiring any topology change. This effectively invalidates Wheeler's argument.

particles, and space-time, made up of *events*. Let us see how these two concepts are defined.

A *particle* – and in general any object – is defined as a conserved entity that has a position and that can move. In fact, the etymology of the word *object* is connected to the latter property. In other words, a particle is a small entity with conserved mass, charge, spin and so on, whose position can vary with time.

Ref. 48 An *event* is a point in space and time. In every physics text, *time* is defined with the help of moving objects, usually called 'clocks', or moving particles, such as those emitted by light sources. Similarly, *length* is defined in terms of objects, either with an old-fashioned ruler or in terms of the motion of light, which is itself motion of particles.

Modern physics has sharpened our definitions of particles and space-time. Quantum mechanics assumes that space-time is given (as a symmetry of the Hamiltonian), and studies the properties of particles and their motion, both for matter and for radiation. Quantum theory has deduced the full list of properties that define a particle. General relativity, and especially cosmology, takes the opposite approach: it assumes that the properties of matter and radiation are given (for example, via their equations of state), and describes in detail the space-time that follows from them, in particular its curvature.

However, one fact remains unchanged throughout all these advances: in the millennium description of nature, *the two concepts of particle and of space-time are each defined with the help of the other*. This circular definition is the origin of the contradictions between quantum mechanics and general relativity. In order to eliminate the contradictions and to formulate a complete theory, we must eliminate this circular definition.

The domain of contradictions: Planck scales

Despite their contradictions and the underlying circular definition, both general relativity and quantum theory are successful theories for the description of nature: they agree with all data. How can this be?

Each theory of modern physics provides a criterion for determining when it is necessary and when classical Galilean physics is no longer applicable. These criteria are the basis for many arguments in the following chapters.

General relativity shows that it is *necessary* to take into account the curvature of empty space* and space-time whenever we approach an object of mass m to within a distance of the order of the Schwarzschild radius r_S, given by

$$r_S = 2Gm/c^2 \ . \tag{72}$$

The gravitational constant G and the speed of light c act as conversion constants. Indeed, as the Schwarzschild radius of an object is approached, the difference between general relativity and the classical $1/r^2$ description of gravity becomes larger and larger. For example, the barely measurable gravitational deflection of light by the Sun is due to the Ref. 42, Ref. 49 light approaching the Sun to within $2.4 \cdot 10^5$ times its Schwarzschild radius. Usually, we are forced to stay away from objects at a distance that is an even larger multiple of the Schwarzschild radius, as shown in Table 2. Only for this reason is general relativity unne-

* In the following, we use the terms 'vacuum' and 'empty space' interchangeably.

TABLE 2 The size, Schwarzschild radius and Compton wavelength of some objects appearing in nature. The lengths in quotation marks make no physical sense, as explained in the text.

Object	Diameter d	Mass m	Schwarzschild radius r_S	Ratio d/r_S	Compton wavelength λ_C (red.)	Ratio d/λ_C
galaxy	≈ 1 Zm	$\approx 5 \cdot 10^{40}$ kg	≈ 70 Tm	$\approx 10^{7}$	'$\approx 10^{-83}$ m'	$\approx 10^{104}$
neutron star	10 km	$2.8 \cdot 10^{30}$ kg	4.2 km	2.4	'$1.3 \cdot 10^{-73}$ m'	$8.0 \cdot 10^{76}$
Sun	1.4 Gm	$2.0 \cdot 10^{30}$ kg	3.0 km	$4.8 \cdot 10^{5}$	'$1.0 \cdot 10^{-73}$ m'	$8.0 \cdot 10^{81}$
Earth	13 Mm	$6.0 \cdot 10^{24}$ kg	8.9 mm	$1.4 \cdot 10^{9}$	'$5.8 \cdot 10^{-68}$ m'	$2.2 \cdot 10^{74}$
human	1.8 m	75 kg	0.11 ym	$1.6 \cdot 10^{25}$	'$4.7 \cdot 10^{-45}$ m'	$3.8 \cdot 10^{44}$
molecule	10 nm	0.57 zg	'$8.5 \cdot 10^{-52}$ m'	$1.2 \cdot 10^{43}$	$6.2 \cdot 10^{-19}$ m	$1.6 \cdot 10^{10}$
atom (^{12}C)	0.6 nm	20 yg	'$3.0 \cdot 10^{-53}$ m'	$2.0 \cdot 10^{43}$	$1.8 \cdot 10^{-17}$ m	$3.2 \cdot 10^{7}$
proton p	2 fm	1.7 yg	'$2.5 \cdot 10^{-54}$ m'	$8.0 \cdot 10^{38}$	$2.0 \cdot 10^{-16}$ m	9.6
pion π	2 fm	0.24 yg	'$3.6 \cdot 10^{-55}$ m'	$5.6 \cdot 10^{39}$	$1.5 \cdot 10^{-15}$ m	1.4
up-quark u	< 0.1 fm	$5 \cdot 10^{-30}$ kg	'$7 \cdot 10^{-57}$ m'	$< 1 \cdot 10^{40}$	$7 \cdot 10^{-14}$ m	< 0.001
electron e	< 4 am	$9.1 \cdot 10^{-31}$ kg	'$1.4 \cdot 10^{-57}$ m'	$< 3 \cdot 10^{39}$	$3.9 \cdot 10^{-13}$ m	$< 1 \cdot 10^{-5}$
neutrino ν_e	< 4 am	$< 3 \cdot 10^{-36}$ kg	'$< 5 \cdot 10^{-63}$ m'	n.a.	$> 1 \cdot 10^{-7}$ m	$< 3 \cdot 10^{-11}$

cessary in everyday life. We recall that objects whose size is given by their Schwarzschild radius are black holes; smaller objects cannot exist. Challenge 33 e

Similarly, quantum mechanics shows that Galilean physics must be abandoned and quantum effects *must* be taken into account whenever an object is approached to within distances of the order of the (reduced) Compton wavelength λ_C, given by

$$\bar{\lambda}_C = \frac{\hbar}{m c} \,. \qquad (73)$$

In this case, Planck's constant $\hbar$ and the speed of light c act as conversion factors to transform the mass m into a length scale. Of course, this length is only relevant if the object is smaller than its own Compton wavelength. At these scales we get relativistic quantum effects, such as particle–antiparticle pair creation or annihilation. Table 2 shows that the approach distance is near to or smaller than the Compton wavelength only in the microscopic world, so that such effects are not observed in everyday life. Only for this reason we do not need quantum field theory to describe common observations.

Combining concepts of quantum field theory and general relativity is required in situations where both conditions are satisfied simultaneously. The necessary approach distance for such situations is calculated by setting $r_S = 2\lambda_C$ (the factor 2 is introduced for simplicity). We find that this is the case when lengths or times are – within a factor of order 1 – of the order of

$$\begin{aligned} l_{Pl} = \sqrt{\hbar G/c^3} &= 1.6 \cdot 10^{-35}\,\text{m, the } \textit{Planck length}, \\ t_{Pl} = \sqrt{\hbar G/c^5} &= 5.4 \cdot 10^{-44}\,\text{s, the } \textit{Planck time}. \end{aligned} \qquad (74)$$

Whenever we approach objects at these scales, both general relativity and quantum mechanics play a role, and effects of *quantum gravity* appear. Because the values of the Planck dimensions are extremely small, this level of sophistication is unnecessary in everyday life, in astronomy and even in particle physics.

In the millennium description of nature, all the contradictions and also the circular definition just mentioned are effective only at Planck scales. You can check this yourself. This is the reason that general relativity and quantum theory work so well in practice.

Challenge 34 e

However, to answer the questions posed at the beginning – why do we live in three dimensions, why are there three interactions, and why is the proton 1836.15 times heavier than the electron? – we require a precise and complete description of nature. To answer these questions, we must understand physics at Planck scales.

In summary, general relativity and quantum theory do contradict each other. However, the domains where these contradictions play a role, the Planck scales, are not accessible by experiment. As a consequence, the contradictions and our lack of knowledge of how nature behaves at the Planck scales have only one effect: we do not see the solutions to the millennium issues.

We note that some researchers argue that the Planck scales specify only one of several domains of nature where quantum mechanics and general relativity apply simultaneously. They mention horizons and the big bang as separate domains. However, it is more appropriate to argue that horizons and the big bang are situations where Planck scales are essential.

Resolving the contradictions

The contradictions between general relativity and quantum theory have little practical consequences. Therefore, for a long time, the contradictions were accommodated by keeping the two theories separate. It is often said that quantum mechanics is valid at small scales and general relativity is valid at large scales. This attitude is acceptable as long as we remain far from the Planck length. However, this accommodating attitude also prevents us from resolving the circular definition, the contradictions and therefore, the millennium issues.

The situation resembles the well-known drawing, Figure 2, by Maurits Escher (b. 1898 Leeuwarden, d. 1972 Hilversum) in which two hands, each holding a pencil, seem to be drawing each other. If one hand is taken as a symbol of vacuum and the other as a symbol of particles, with the act of drawing taken as the act of defining, the picture gives a description of twentieth-century physics. The apparent circular definition is solved by recognizing that the two concepts (the two hands) both originate from a third, hidden concept. In the picture, this third entity is the hand of the artist. In physics, the third concept is the common origin of vacuum and particles.

We thus conclude that the contradictions in physics and the circular definition are solved by *common constituents* for vacuum and matter. In order to find out what these common constituents are and what they are not, we must explore the behaviour of nature at the Planck scales.

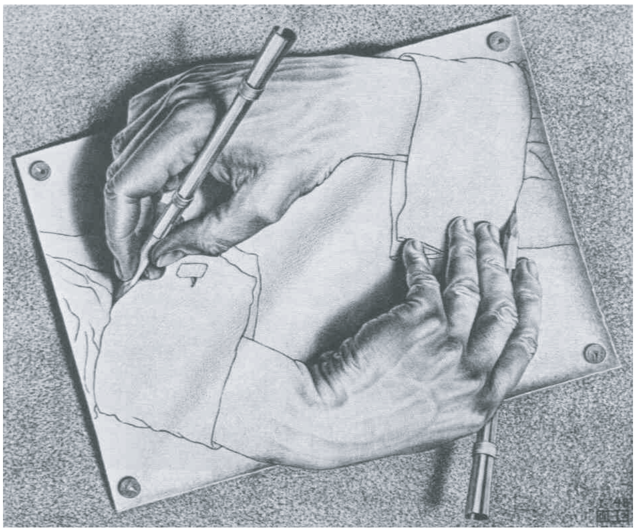


FIGURE 2 'Tekenen' by Maurits Escher, 1948 – a metaphor for the way in which 'particles' and 'space-time' are defined: each with the help of the other (© M.C. Escher Heirs).

The origin of points

General relativity is built on the *assumption* that space is a continuum of points. Already at school we learn that lines, surfaces and areas are made of points. We take this as granted, because we imagine that finer and finer measurements are always possible. And all experiments so far agree with the assumption. Fact is: in this reasoning, we first idealized measurement rulers – which are made of matter – and then 'deduced' that points in space exist.

Quantum theory is built on the *assumption* that elementary particles are point-like. We take this as granted, because we imagine that collisions at higher and higher energy are possible that allow elementary particles to get as close as possible. And all experiments so far agree with the assumption. Fact is: in this reasoning, we first imagined infinite energy and momentum values – which is a statement on time and space properties – and then 'deduce' that point particles exist.

Page 58

▷ The use of points in space and of separate, point-like particles are the reasons for the mistaken vacuum energy calculation (71) that is wrong by 120 orders of magnitude.

In short, only the circular definition of space and matter allows us to define points and point particles. This puts us in a strange situation. On the one hand, experiment tells us that describing nature with space points and with point particles works. On the other hand, reason tells us that this is a fallacy and cannot be correct at Planck scales. We need a solution.

SUMMARY ON THE CLASH BETWEEN THE TWO THEORIES

General relativity and quantum theory contradict each other. In practice however, this happens only at Planck scales; and this includes horizons. The reason for the contradiction is our insistence on a circular definition of space and particles. Indeed, we *need* this circularity: Only such a circular definition allows us to define *space points* and *point particles* at all.

In order to solve the contradictions between general relativity and quantum theory and in order to understand nature at Planck scales, we must introduce *common* constituents for space and particles. But common constituents have an important consequence: common constituents force us to stop using points to describe nature. We now explore this connection.

Chapter 4

DOES MATTER DIFFER FROM VACUUM?

Vol. II, page 24

The appearance of the quantum of action in the description of motion leads to limitations for all measurements: Heisenberg's indeterminacy relations. These relations, when combined with the effects of gravitation, imply an almost unbelievable series of consequences for the behaviour of nature at Planck scales. The most important ones are the necessity to abandon points, instants and events, and the equivalence of vacuum and matter. Here we show how these surprising and important conclusions follow from simple arguments based on the indeterminacy relations, the Compton wavelength and the Schwarzschild radius.

Farewell to instants of time

> Time is composed of time atoms ... which in fact are indivisible.
>
> Maimonides**

Measurement limits appear most clearly when we investigate the properties of clocks and metre rules. Is it possible to construct a clock that is able to measure time intervals shorter than the Planck time? Surprisingly, the answer is no, even though the time–energy indeterminacy relation $\Delta E \Delta t \geqslant \hbar$ seems to indicate that by making ΔE large enough, we can make Δt arbitrary small. Ref. 50, Ref. 51

Every clock is a device with some moving parts. The moving parts can be mechanical wheels, or particles of matter in motion, or changing electrodynamic fields (i.e., photons), or decaying radioactive particles. For each moving component of a clock the indeterminacy relation applies. As explained most clearly by Michael Raymer, the indeterminacy relation for two non-commuting variables describes two different, but related, situations: it makes a statement about standard *deviations* of *separate* measurements on *many* identical systems; and it describes the measurement *precision* for a *joint* measurement on a *single* system. In what follows, we will consider only the second situation. Ref. 52, Ref. 53 Ref. 54

For a clock to be useful, we need to know both the time and the energy of each hand. Otherwise it would not be a recording device. More generally, a clock must be a classical system. We need the combined knowledge of the non-commuting variables for each moving component of the clock. Let us focus on the component with the largest time indeterminacy Δt. It is evident that the smallest time interval δt that can be measured by

** Moses Maimonides (b. 1135 Cordoba, d. 1204 Egypt) was a physician, philosopher and influential theologian. However, there is no evidence for 'time atoms' in nature, as explained below. Page 75

a clock is always larger than the quantum limit, i.e., larger than the time indeterminacy Δt for the most 'uncertain' component. Thus we have

$$\delta t \geqslant \Delta t \geqslant \frac{\hbar}{\Delta E} , \qquad (75)$$

where ΔE is the energy indeterminacy of the moving component. Now, ΔE must be smaller than the total energy $E = c^2 m$ of the component itself: $\Delta E < c^2 m$.* Furthermore, a clock provides information, so signals have to be able to leave it. Therefore the clock must *not* be a black hole: its mass m must be smaller than a black hole of its size, i.e., $m \leqslant c^2 l/G$, where l is the size of the clock (neglecting factors of order unity). Finally, for a sensible measurement of the time interval δt, the size l of the clock must be smaller than $c\,\delta t$, because otherwise different parts of the clock could not work together to produce the same time display: $l < c\delta t$.** If we combine these three conditions, we get

$$\delta t \geqslant \frac{\hbar G}{c^5 \delta t} \qquad (76)$$

or

$$\delta t \geqslant \sqrt{\frac{\hbar G}{c^5}} = t_{\text{Pl}} . \qquad (77)$$

In summary, from three simple properties of any clock – namely, that it is only a single clock, that we can read its dial, and that it gives sensible read-outs – we conclude that *clocks cannot measure time intervals shorter than the Planck time*. Note that this argument is independent of the nature of the clock mechanism. Whether the clock operates by gravitational, electrical, mechanical or even nuclear means, the limit still applies.***

Ref. 59 The same conclusion can be reached in other ways. For example, any clock small enough to measure small time intervals necessarily has a certain energy indeterminacy due to the indeterminacy relation. Meanwhile, on the basis of general relativity, any energy density induces a deformation of space-time, and signals from the deformed region arrive with a certain delay due to that deformation. The energy indeterminacy of the source leads to an indeterminacy in the deformation, and thus in the delay. The expression from general relativity for the deformation of the time part of the line element due to a mass m is $\delta t = mG/lc^3$. From the mass–energy relation, we see that an energy spread

Vol. II, page 163 Ref. 49

* Physically, this condition means being sure that there is only *one* clock: if $\Delta E > E$, it would be impossible to distinguish between a single clock and a clock–anticlock pair created from the vacuum, or a component together with two such pairs, and so on.

** It is amusing to explore how a clock *larger* than $c\,\delta t$ would stop working, as a result of the loss of rigidity in its components. Challenge 35 s

*** Gravitation is essential here. The present argument differs from the well-known study on the limitations of clocks due to their mass and their measuring time which was published by Salecker and Wigner and summarized in pedagogical form by Zimmerman. In our case, both quantum mechanics and gravity are included, and therefore a different, lower, and more fundamental limit is found. Also the discovery of black hole radiation does not change the argument: black hole radiation notwithstanding, measurement devices cannot exist inside black holes. Ref. 55, Ref. 56, Ref. 57, Ref. 58

ΔE produces an indeterminacy Δt in the delay:

$$\Delta t = \frac{\Delta E\, G}{l\, c^5} \,. \tag{78}$$

This determines the precision of the clock. Furthermore, the energy indeterminacy of the clock is fixed by the indeterminacy relation for time and energy $\Delta E \geqslant \hbar/\Delta t$. Combining all this, we again find the relation $\delta t \geqslant t_{\mathrm{Pl}}$ for the minimum measurable time.

We are forced to conclude that in nature, it is impossible to measure time intervals shorter than the Planck time. Thus

▷ In nature there is a minimum time interval.

In other words, *at Planck scales the term 'instant of time' has no theoretical or experimental basis.* But let us go on. Special relativity, quantum mechanics and general relativity all rely on the idea that time can be defined for all points of a given reference frame. However, two clocks a distance l apart cannot be synchronized with arbitrary precision. Since the distance between two clocks cannot be measured with an error smaller than the Planck length l_{Pl}, and transmission of signals is necessary for synchronization, it is not possible to synchronize two clocks with a better precision than $l_{\mathrm{Pl}}/c = t_{\mathrm{Pl}}$, the Planck time. So use of a single *time coordinate* for a whole reference frame is only an approximation. *Reference frames do not have a single time coordinate at Planck scales.*

Moreover, since the time difference between events can only be measured within a Planck time, for two events distant in time by this order of magnitude, it is not possible
Ref. 60 to say with complete certainty which of the two precedes the other. But if events cannot be ordered, then the very concept of time, which was introduced into physics to describe sequences, makes no sense at Planck scales. In other words, after dropping the idea of a common time coordinate for a complete frame of reference, we are forced to drop the idea of time at a single 'point' as well. *The concept of 'proper time' loses its meaning at Planck scales.*

Farewell to points in space

> "Our greatest pretenses are built up not to hide the evil and the ugly in us, but our emptiness. The hardest thing to hide is something that is not there."
> Eric Hoffer,* *The Passionate State of Mind*

In a similar way, we can deduce that it is impossible to make a metre rule, or any other
Ref. 24 length-measuring device, that is able to measure lengths shorter than the Planck length. Obviously, we can already deduce this from $l_{\mathrm{Pl}} = c\, t_{\mathrm{Pl}}$, but an independent proof is also possible.

For any length measurement, joint measurements of position and momentum are necessary. The most straightforward way to measure the distance between two points is to put an object at rest at each position. Now, the minimal length δl that can be measured

* Eric Hoffer (b. 1902 New York City, d. 1983 San Francisco), philosopher.

must be larger than the position indeterminacy of the two objects. From the indeterminacy relation we know that neither object's position can be determined with a precision Δl better than that given by $\Delta l\,\Delta p = \hbar$, where Δp is the momentum indeterminacy. The requirement that there be only one object at each end (avoiding pair production from the vacuum) means that $\Delta p < mc$: together, these requirements give

$$\delta l \geqslant \Delta l \geqslant \frac{\hbar}{mc} \,. \tag{79}$$

Furthermore, the measurement cannot be performed if signals cannot leave the objects; thus, they cannot be black holes. Therefore their masses must be small enough for their Schwarzschild radius $r_S = 2Gm/c^2$ to be less than the distance δl separating them. Again omitting the factor of 2, we get

$$\delta l \geqslant \sqrt{\frac{\hbar G}{c^3}} = l_{Pl} \,. \tag{80}$$

Length measurements are limited by the Planck length.

Another way to deduce this limit reverses the roles of general relativity and quantum theory. To measure the distance between two objects, we have to localize the first object with respect to the other within a certain interval Δx. The corresponding energy indeterminacy obeys $\Delta E = c(c^2m^2 + (\Delta p)^2)^{1/2} \geqslant c\hbar/\Delta x$. However, general relativity shows that a small volume filled with energy changes the curvature of space-time, and thus changes the metric of the surrounding space. For the resulting distance change Δl, compared with empty space, we find the expression $\Delta l \approx G\Delta E/c^4$. In short, if we localize the first particle in space with a precision Δx, the distance to a second particle is known only with precision Δl. The minimum length δl that can be measured is obviously larger than either of these quantities; inserting the expression for ΔE, we find again that the minimum measurable length δl is given by the Planck length.

Ref. 42, Ref. 49
Ref. 24
Ref. 61
Ref. 62, Ref. 63
Ref. 26

We note that every length measurement requires a joint measurement of position and momentum. This is particularly obvious if we approach a metre ruler to an object, but it is equally true for any other length measurement.

We note that, since the Planck length is the shortest possible length, there can be no observations of quantum-mechanical effects for a situation where the corresponding de Broglie or Compton wavelength is smaller than the Planck length. In proton–proton collisions we observe both pair production and interference effects. In contrast, the Planck limit implies that in everyday, macroscopic situations, such as car–car collisions, we cannot observe embryo–antiembryo pair production and quantum interference effects.

Another way to convince oneself that points have no meaning is to observe that a point is an entity with vanishing volume; however, the minimum volume possible in nature is the Planck volume $V_{Pl} = l_{Pl}^3$.

We conclude that the Planck units not only provide *natural* units; they also provide – within a factor of order one – the *limit* values of space and time intervals.

In summary, from two simple properties common to all length-measuring devices, namely that they are discrete and that they can be read, we arrive at the conclusion that

▷ Lengths smaller than the Planck length cannot be measured.

Whatever method is used, be it a metre rule or time-of-flight measurement, we cannot overcome this fundamental limit. It follows that *the concept of a 'point in space' has no experimental or theoretical basis*. In other terms,

▷ In nature there is a minimum length interval.

The limitations on length measurements imply that we cannot speak of continuous space, except in an approximate sense. As a result of the lack of measurement precision at Planck scales, the concepts of spatial order, of translation invariance, of isotropy of the vacuum and of global coordinate systems have no experimental basis.

The generalized indeterminacy relation

Ref. 24 The limit values for length and time measurements are often expressed by the so-called *generalized indeterminacy relation*

$$\Delta p \Delta x \geqslant \hbar/2 + f \frac{G}{c^3} (\Delta p)^2 \qquad (81)$$

or

$$\Delta p \Delta x \geqslant \hbar/2 + f \frac{l_{\mathrm{Pl}}^2}{\hbar} (\Delta p)^2 , \qquad (82)$$

where f is a numerical factor of order unity. A similar expression holds for the time–energy indeterminacy relation. The first term on the right-hand side is the usual quantum-mechanical indeterminacy. The second term is negligible for everyday energies, and is significant only near Planck energies; it is due to the changes in space-time induced by gravity at these high energies. You should be able to show that the generalized principle (81) implies that Δx can never be smaller than $f^{1/2} l_{\mathrm{Pl}}$. Challenge 36 e

The generalized indeterminacy relation is derived in exactly the same way in which Heisenberg derived the original indeterminacy relation $\Delta p \Delta x \geqslant \hbar/2$, namely by studying the scattering of light by an object under a microscope. A careful re-evaluation of the process, this time including gravity, yields equation (81). For this reason, *all* descriptions that unify quantum mechanics and gravity must yield this relation, and indeed all known approaches do so. Ref. 24, Ref. 64, Ref. 65, Ref. 66, Ref. 67, Ref. 68

Farewell to space-time continuity

> Ich betrachte es als durchaus möglich, dass die Physik nicht auf dem Feldbegriff begründet werden kann, d.h. auf kontinuierlichen Gebilden. Dann bleibt von meinem ganzen Luftschloss inklusive Gravitationstheorie nichts bestehen.*
> Albert Einstein, 1954, in a letter to Michele Besso.

* 'I consider it as quite possible that physics cannot be based on the field concept, i.e., on continuous structures. In that case, nothing remains of my castle in the air, gravitation theory included.'

The classical description of nature is based on continuity: it involves and allows differences of time and space that are as small as can be imagined. Between any two points in time or space, the existence of infinitely many other points is assumed. Measurement results of arbitrary small values are deemed possible. The same is valid for action values.

However, quantum mechanics begins with the realization that the classical concept of action makes no sense below the value of $\hbar/2$; similarly, unified theories begin with the realization that the classical concepts of time and length make no sense below Planck scales. Therefore, *the continuum description of space-time has to be abandoned* in favour of a more appropriate description.

The minimum length distance, the minimum time interval, and equivalently, the new, generalized indeterminacy relation appearing at Planck scales show that space, time and in particular, space-time, are not well described as a continuum. Inserting $c\Delta p \geqslant \Delta E \geqslant \hbar/\Delta t$ into equation (81), we get

Ref. 69

$$\Delta x \Delta t \geqslant \hbar G/c^4 = t_{\mathrm{Pl}} l_{\mathrm{Pl}} , \tag{83}$$

which of course has no counterpart in standard quantum mechanics. This shows that also space-time *events do not exist*. The concept of an 'event', being a combination of a 'point in space' and an 'instant of time', loses its meaning for the description of nature at Planck scales.

Interestingly, the view that continuity must be abandoned is almost one hundred years old. Already in 1917, Albert Einstein wrote in a letter to Werner Dällenbach:

Ref. 70

> Wenn die molekulare Auffassung der Materie die richtige (zweckmässige) ist, d.h. wenn ein Teil Welt durch eine endliche Zahl bewegter Punkte darzustellen ist, so enthält das Kontinuum der heutigen Theorie *zu viel* Mannigfaltigkeit der Möglichkeiten. Auch ich glaube, dass dieses zu viel daran schuld ist, dass unsere heutige Mittel der Beschreibung an der Quantentheorie scheitern. Die Frage scheint mir, wie man über ein Diskontinuum Aussagen formulieren kann, ohne ein Kontinuum (Raum-Zeit) zu Hilfe zu nehmen; letzteres wäre als eine im Wesen des Problems nicht gerechtfertigte zusätzliche Konstruktion, der nichts "Reales" entspricht, aus der Theorie zu verbannen. Dazu fehlt uns aber leider noch die mathematische Form. Wie viel habe ich mich in diesem Sinne schon geplagt!
>
> Allerdings sehe ich auch hier prinzipielle Schwierigkeiten. Die Elektronen (als Punkte) wären in einem solchen System letzte Gegebenheiten (Bausteine). Gibt es überhaupt letzte Bausteine? Warum sind diese alle von gleicher Grösse? Ist es befriedigend zu sagen: Gott hat sie in seiner Weisheit alle gleich gross gemacht, jedes wie jedes andere, weil er so wollte; er hätte sie auch, wenn es ihm gepasst hätte, verschieden machen können. Da ist man bei der Kontinuum-Auffassung besser dran, weil man nicht von Anfang an die Elementar-Bausteine angeben muss. Ferner die alte Frage vom Vakuum! Aber diese Bedenken müssen verblassen hinter der blendenden Tatsache: Das Kontinuum ist ausführlicher als die zu beschreibenden Dinge...
>
> Lieber Dällenbach! Was hilft alles Argumentieren, wenn man nicht bis zu einer befriedigenden Auffassung durchdringt; das aber ist verteufelt schwer.

> Es wird einen schweren Kampf kosten, bis man diesen Schritt, der uns da vorschwebt, wirklich gemacht haben wird. Also strengen Sie Ihr Gehirn an, vielleicht zwingen Sie es.*

The second half of this text will propose a way to rise to the challenge. At this point however, we first complete the exploration of the limitations of continuum physics.

In 20th century physics, space-time points are idealizations of events – but this idealization is inadequate. The use of the concept of 'point' is similar to the use of the concept of 'aether' a century ago: it is impossible to measure or detect.

▷ Like the 'aether', also 'points' lead reason astray.

All paradoxes resulting from the infinite divisibility of space and time, such as Zeno's argument on the impossibility of distinguishing motion from rest, or the Banach–Tarski paradox, are now avoided. We can dismiss them straight away because of their incorrect premises concerning the nature of space and time.

The consequences of the Planck limits for measurements of time and space can be expressed in other ways. It is often said that given any two points in space or any two instants of time, there is always a third in between. Physicists sloppily call this property continuity, while mathematicians call it denseness. However, at Planck scales this property cannot hold, since there are no intervals smaller than the Planck time. Thus points and instants are not dense, and

▷ Between two points there is not always a third.

This results again means that *space and time are not continuous.* Of course, at large scales they are – approximately – continuous, in the same way that a piece of rubber or a liquid seems continuous at everyday scales, even though it is not at a small scale. But in nature, space, time and space-time are not continuous entities.

* 'If the molecular conception of matter is the right (appropriate) one, i.e., if a part of the world is to be represented by a finite number of moving points, then the continuum of the present theory contains *too great* a manifold of possibilities. I also believe that this 'too great' is responsible for our present means of description failing for quantum theory. The questions seems to me how one can formulate statements about a discontinuum without using a continuum (space-time) as an aid; the latter should be banned from the theory as a supplementary construction not justified by the essence of the problem, which corresponds to nothing "real". But unfortunately we still lack the mathematical form. How much have I already plagued myself in this direction!

Yet I also see difficulties of principle. In such a system the electrons (as points) would be the ultimate entities (building blocks). Do ultimate building blocks really exist? Why are they all of equal size? Is it satisfactory to say: God in his wisdom made them all equally big, each like every other one, because he wanted it that way; he could also have made them, if he had wanted, all different. With the continuum viewpoint one is better off, because one doesn't have to prescribe elementary building blocks from the outset. Furthermore, the old question of the vacuum! But these considerations must pale beside the dazzling fact: The continuum is more ample than the things to be described...

Dear Dällenbach! All arguing does not help if one does not achieve a satisfying conception; but this is devilishly difficult. It will cost a difficult fight until the step that we are thinking of will be realized. Thus, squeeze your brain, maybe you can force it.'

Page 87 Compare this letter to what Einstein wrote almost twenty and almost forty years later.

But there is more to come. The very existence of a minimum length contradicts the theory of special relativity, in which it is shown that lengths undergo Lorentz contraction when the frame of reference is changed. There is only one conclusion: special relativity (and general relativity) cannot be correct at very small distances. Thus,

▷ Space-time is not Lorentz-invariant (nor diffeomorphism-invariant) at Planck scales.

All the symmetries that are at the basis of special and general relativity are only approximately valid at Planck scales.

The imprecision of measurement implies that most familiar concepts used to describe spatial relations become useless. For example, the concept of a *metric* loses its usefulness at Planck scales, since distances cannot be measured with precision. So it is impossible to say whether space is flat or curved. The impossibility of measuring lengths exactly is equivalent to fluctuations of the curvature, and thus of gravity.

Ref. 24, Ref. 71

In short, space and space-time are not smooth at Planck scales. This conclusion has important implications. For example, the conclusion implies that certain mathematical solutions found in books on general relativity, such as the Eddington–Finkelstein coordinates and the Kruskal–Szekeres coordinates do *not* describe nature! Indeed, these coordinate systems, which claim to show that space-time goes on *behind* the horizon of a black hole, are based on the idea that space-time is smooth everywhere. However, quantum physics shows that space-time is not smooth at the horizon, but fluctuates wildly there. In short, quantum physics confirms what common sense already knew: *Behind a horizon, nothing can be observed, and thus there is nothing there.*

Farewell to dimensionality

Even the number of spatial dimensions makes no sense at Planck scales. Let us remind ourselves how to determine this number experimentally. One possible way is to determine how many points we can choose in space such that all the distances between them are equal. If we can find at most n such points, the space has $n-1$ dimensions. But if reliable length measurement at Planck scales is not possible, there is no way to determine reliably the number of dimensions of space with this method.

Another way to check for three spatial dimensions is to make a knot in a shoe string and glue the ends together: since it stays knotted, we know that space has three dimensions, because there is a mathematical theorem that in spaces with greater or fewer than three dimensions, knots do not exist. Again, at Planck scales, we cannot say whether a string is knotted or not, because measurement limits at crossings make it impossible to say which strand lies above the other.

There are many other methods for determining the dimensionality of space.* In all cases, the definition of dimensionality is based on a precise definition of the concept of

* For example, we can determine the dimension using only the topological properties of space. If we draw a so-called *covering* of a topological space with open sets, there are always points that are elements of several sets of the covering. Let p be the maximal number of sets of which a point can be an element in a given covering. The minimum value of p over all possible coverings, minus one, gives the dimension of the space.

In fact, if physical space is not a manifold, the various methods for determining the dimensionality may give different answers. Indeed, for linear spaces without norm, the dimensionality cannot be defined in a neighbourhood. At Planck scales, however, length measurements do not allow us to say whether a given point is inside or outside a given region. In short, whatever method we use, the lack of precise length measurements means that

▷ At Planck scales, the dimensionality of physical space is not defined.

Farewell to the space-time manifold

The reasons for the problems with space-time become most evident when we remember Ref. 72 Euclid's well-known definition: 'A point is that which has no part.' As Euclid clearly understood, a *physical* point, as an idealization of position, cannot be defined without some measurement method. *Mathematical* points, however, can be defined without reference to a metric. They are just elements of a set, usually called a 'space'. (A 'measurable' or 'metric' space is a set of points equipped with a measure or a metric.)

In the case of physical space-time, the concepts of measure and of metric are more fundamental than that of a point. Confusion between physical and mathematical space and points arises from the failure to distinguish a mathematical metric from a physical length measurement.*

Vol. I, page 57 Ref. 73 Perhaps the most beautiful way to make this point is the Banach–Tarski theorem, which clearly shows the limits of the concept of *volume*. The theorem states that a sphere made up of *mathematical points* can be cut into five pieces in such a way that the pieces can be put together to form two spheres, each of the same volume as the original one. However, the necessary 'cuts' are infinitely curved and detailed: the pieces are wildly disconnected. For physical matter such as gold, unfortunately – or fortunately – the existence of a minimum length, namely the atomic distance, makes it impossible to perform such a cut. For vacuum, the puzzle reappears. For example, the energy of zero-point fluctuations is given by the density times the volume; following the Banach–Tarski theorem, the zero-point energy content of a single sphere should be equal to the zero-point energy of two similar spheres each of the same volume as the original one. The paradox is resolved by the Planck length, which provides a fundamental length scale even for vacuum, thus making infinitely complex cuts impossible. Therefore, the concept of volume is only well defined at Planck scales if a minimum length is introduced.

To sum up:

▷ Physical space-time cannot be a set of mathematical points.

But there are more surprises. At Planck scales, since both temporal and spatial order break down, there is no way to say if the distance between two nearby space-time regions

---

unique way. Different definitions (fractal dimension, Lyapunov dimension, etc.) are possible.

* Where does the incorrect idea of continuous space-time have its roots? In everyday life, as well as in physics, space-time is a book-keeping device introduced to describe observations. Its properties are extracted from the properties of observables. Since observables can be added and multiplied, like numbers, we infer that they can take continuous values, and, in particular, arbitrarily small values. It is then possible to define points and sets of points. A special field of mathematics, topology, shows how to start from a set of points and construct, with the help of neighbourhood relations and separation properties, first a *topological space*, then, with the help of a metric, a *metric space*. With the appropriate compactness and connectedness relations, a *manifold*, characterized by its dimension, metric and topology, can be constructed.

is space-like or time-like.

▷ At Planck scales, time and space cannot be distinguished from each other.

In addition, we cannot state that the topology of space-time is fixed, as general relativity implies. The topology changes, mentioned above, that are required for particle reactions do become possible. In this way another of the contradictions between general relativity and quantum theory is resolved.

In summary, space-time at Planck scales is not continuous, not ordered, not endowed with a metric, not four-dimensional, and not made up of points. It satisfies none of the defining properties of a manifold.* We conclude that *the concept of a space-time manifold has no justification at Planck scales.* This is a strong result. Even though both general relativity and quantum mechanics use continuous space-time, the combined theory does not.

Farewell to observables, symmetries and measurements

If space and time are not continuous, no quantities defined as derivatives with respect to space or time are precisely defined. Velocity, acceleration, momentum, energy and so on are only well defined under the assumption of continuity. That important tool, the evolution equation, is based on derivatives and can thus no longer be used. Therefore the Schrödinger and Dirac equations lose their basis. Concepts such as 'derivative', 'divergence-free' and 'source free' lose their meaning at Planck scales.

All physical observables are defined using length and time measurements. Each physical unit is a product of powers of length and time (and mass) units. (In the SI system, electrical quantities have a separate base quantity, the ampere, but the argument still holds: the ampere is itself defined in terms of a force, which is measured using the three base units of length, time and mass.) Since time and length are not continuous, *at Planck scales, observables cannot be described by real numbers.*

In addition, if time and space are not continuous, the usual expression for an observable field, $A(t, x)$, does not make sense: we have to find a more appropriate description. *Physical fields cannot exist at Planck scales.* Quantum mechanics also relies on the possibility to add wave functions; this is sometimes called the *superposition principle*. Without fields and superpositions, all of quantum mechanics comes crumbling down.

The lack of real numbers has severe consequences. It makes no sense to define multiplication of observables by real numbers, but only by a discrete set of numbers. Among other implications, this means that observables do not form a linear algebra. *Observables are not described by operators at Planck scales.* In particular, the most important observables are the gauge potentials. Since they do not form an algebra, *gauge symmetry is not valid at Planck scales.* Even innocuous-looking expressions such as $[x_i, x_j] = 0$ for $x_i \neq x_j$, which are at the root of quantum field theory, become meaningless at Planck scales. Since at those scales superpositions cannot be backed up by experiment, even the famous Wheeler–DeWitt equation, sometimes assumed to describe quantum gravity, cannot be valid.

Vol. V, page 359

* A manifold is what looks *locally* like a Euclidean space. The exact definition can be found in the previous volume.

Similarly, permutation symmetry is based on the premise that we can distinguish two points by their coordinates, and then exchange particles between those locations. As we have just seen, this is not possible if the distance between the two particles is very small. We conclude that *permutation symmetry has no experimental basis at Planck scales.*

Even discrete symmetries, like charge conjugation, space inversion and time reversal, cannot be correct in this domain, because there is no way to verify them exactly by measurement. *CPT symmetry is not valid at Planck scales.*

Finally we note that all types of scaling relations break down at small scales, because of the existence of a smallest length. As a result, the *renormalization group breaks down at Planck scales.*

In summary, due to the impossibility of accurate measurements,

▷ All symmetries break down at Planck scales.

For example, supersymmetry cannot be valid at Planck scale. All mentioned conclusions are consistent: if there are no symmetries at Planck scales, there are also no observables, since physical observables are representations of symmetry groups. And thus,

▷ The concept of measurement has no significance at Planck scales.

This results from the limitations on time and length measurements.

Can space or space-time be a lattice?

Let us take a breath. Can a space or even a space-time lattice be an alternative to continuity?

Ref. 74 Discrete models of space-time have been studied since the 1940s. Recently, the idea
Ref. 75 that space or space-time could be described as a lattice – like a crystal – has been explored
Ref. 76 most notably by David Finkelstein and by Gerard 't Hooft. The idea of space as a lattice is based on the idea that, if there is a minimum distance, then all distances are multiples
Ref. 77 of this minimum.

In order to get an isotropic and homogeneous situation for large, everyday scales,
Ref. 78 the structure of space cannot be periodic, but must be *random*. But not only must it be random in space, it must also be *fluctuating in time*. In fact, any fixed structure for space-time would violate the result that there are no lengths smaller than the Planck length: as a result of the Lorentz contraction, any moving observer would find lattice distances smaller than the Planck value. Worse still, the fixed lattice idea conflicts with general relativity, in particular with the diffeomorphism-invariance of the vacuum.

Thus, *neither space nor space-time can be a lattice.* A minimum distance does exist in nature; however, we cannot hope that all other distances are simple multiples of it.

▷ Space is not discrete. Neither is space-time.

We will discover more evidence for this negative conclusion later on.

But in fact, many discrete models of space and time have a much bigger limitation. Any such model has to answer a simple question: Where is a particle *during* the jump from one lattice point to the next? This simple question eliminates most naive space-time models.

A glimpse of quantum geometry

Given that space-time is not a set of points or events, it must be something else. We have three hints at this stage. The first is that in order to improve our description of motion we must abandon 'points', and with them, abandon the *local* description of nature. Both quantum mechanics and general relativity assume that the phrase 'observable at a point' has a precise meaning. Because it is impossible to describe space as a manifold, this expression is no longer useful. The unification of general relativity and quantum physics forces the adoption of a *non-local* description of nature at Planck scales. This is the first hint.

The existence of a minimum length implies that there is no way to physically distinguish between locations that are even closer together. We are tempted to conclude that *no* pair of locations can be distinguished, even if they are one metre apart, since on any path joining two points, no two locations that are close together can be distinguished. The problem is similar to the question about the size of a cloud or of an atom. If we measure water density or electron density, we find non-vanishing values at any distance from the centre of the cloud or the atom; however, an effective size can still be defined, because it is very unlikely that the effects of the presence of a cloud or of an atom can be seen at distances much larger than this effective size. Similarly, we can guess that two points in space-time at a macroscopic distance from each other can be distinguished because the probability that they will be confused drops rapidly with increasing distance. In short, we are thus led to a *probabilistic* description of space-time. This is the second hint. Space-time becomes a macroscopic observable, a *statistical* or *thermodynamic limit* of some microscopic entities. This is our second hint.

We note that a fluctuating structure for space-time also avoids the problems of fixed structures with Lorentz invariance. In summary, the experimental observations of special relativity – Lorentz invariance, isotropy and homogeneity – together with the notion of a minimum distance, point towards a description of space-time as *fluctuating*. This is the third hint.

Ref. 27
Several research approaches in quantum gravity have independently confirmed that a *non-local* and *fluctuating* description of space-time at Planck scales resolves the contradictions between general relativity and quantum theory. These are our first results on quantum geometry. To clarify the issue, we turn to the concept of the particle.

Farewell to point particles

In every example of motion, some object is involved. One of the important discoveries of the natural sciences was that all objects are composed of small constituents, called *elementary particles*. Quantum theory shows that all *composite*, non-elementary objects have a finite, non-vanishing size. The naive statement is: a particle is elementary if it behaves like a point particle. At present, only the leptons (electron, muon, tau and the neutrinos), the quarks, the radiation quanta of the electromagnetic, weak and strong nuclear interactions (the photon, the W and Z bosons, and the gluons) and the Higgs boson have been found to be elementary. Protons, atoms, molecules, cheese, people, galaxies and so

Vol. V, page 262

Page 61 on are all composite, as shown in Table 2.

Although the naive definition of 'elementary particle' as point particle is all we need in the following argument, the definition is not precise. It seems to leave open the possibility that future experiments could show that electrons or quarks are not elementary. This is not so! In fact, the precise definition is the following:

Vol. IV, page 107

▷ Any particle is elementary if it is *smaller than its own Compton wavelength*.

If such a small particle were composite, there would be a lighter particle inside it, which would have a larger Compton wavelength than the composite particle. This is impossible, since the size of a composite particle must be larger than the Compton wavelength of its components. (The alternative possibility that all components are heavier than the composite does not lead to satisfying physical properties: for example, it leads to intrinsically unstable components.) Ref. 79

The *size* of an object, such as those given in Table 2, is defined as the length at which differences from point-like behaviour are observed. The size d of an object is determined by measuring how it scatters a beam of probe particles. For example, the radius of the atomic nucleus was determined for the first time in Rutherford's experiment using alpha particle scattering. In daily life as well, when we look at objects, we make use of scattered photons. In general, in order for scattering to be useful, the effective wavelength $\lambda = \hbar/mv$ of the probe must be smaller than the object size d to be determined. We thus need $d > \lambda = \hbar/mv \geqslant \hbar/mc$. In addition, in order for a scattering experiment to be possible, the object must not be a black hole, since, if it were, it would simply swallow the approaching particle. This means that its mass m must be smaller than that of a black hole of the same size; in other words, from equation (72) we must have $m < dc^2/G$. Combining this with the previous condition we get, for the size d of an object, the relation

$$d > \sqrt{\frac{\hbar G}{c^3}} = l_{\mathrm{Pl}} . \qquad (84)$$

In other words, there is no way to observe that an object is smaller than the Planck length. Thus,

▷ There is no way to deduce from observations that a particle is point-like.

The term 'point particle' makes no sense at all.

Of course, there is a relation between the existence of a minimum length for empty space and the existence of a minimum length for objects. If the term 'point of space' is meaningless, then the term 'point particle' is also meaningless. And again, the lower limit on particle size results from the combination of quantum theory and general relativity.*

The minimum size for particles can be tested. A property connected with the size is the electric dipole moment. This describes the deviation of its charge distribution from

* We note that the existence of a minimum size for a particle has nothing to do with the impossibility, in quantum theory, of localizing a particle to within less than its Compton wavelength.

spherical. Some *predictions* from the standard model of elementary particles give as an
Ref. 80 upper limit for the electron dipole moment d_e a value of

$$\frac{|d_e|}{e} < 10^{-39}\,\mathrm{m}\,, \tag{85}$$

where e is the charge of the electron. This predicted value is ten thousand times smaller than the Planck length l_{Pl}. Since the Planck length is the smallest possible length, we
Ref. 81 seem to have a contradiction here. However, a more careful and recent prediction from the standard model only states

$$\frac{|d_e|}{e} < 3 \cdot 10^{-23}\,\mathrm{m}\,, \tag{86}$$

Ref. 82 which is not in contradiction with the minimal length. The experimental limit in 2013 is

$$\frac{|d_e|}{e} < 8.7 \cdot 10^{-31}\,\mathrm{m}\,. \tag{87}$$

In the coming years, the experimental limit value will approach the Planck length. In summary, no point particle is known. In fact, not even a particle smaller than the Planck length is known.

Farewell to particle properties

Planck scales have other strange consequences. In quantum field theory, the difference between a virtual particle and a real particle is that a real particle is 'on shell', obeying $E^2 = m^2c^4 + p^2c^2$, whereas a virtual particle is 'off shell'. Because of the fundamental limits of measurement precision, *at Planck scales we cannot determine whether a particle is real or virtual.*

That is not all. Antimatter can be described as matter moving backwards in time. Since the difference between backwards and forwards cannot be determined at Planck scales, *matter and antimatter cannot be distinguished at Planck scales.*

Every particle is characterized by its spin. Spin describes two properties of a particle: its behaviour under rotations (and thus, if the particle is charged, its behaviour in magnetic fields) and its behaviour under particle exchange. The wave function of a particle with spin 1 remains invariant under a rotation of 2π, whereas that of a particle with spin 1/2 changes sign. Similarly, the combined wave function of two particles with spin 1 does not change sign under exchange of particles, whereas for two particles with spin 1/2 it does.

We see directly that both transformations are impossible to study at Planck scales. Given the limit on position measurements, the position of a rotation axis cannot be well defined, and rotations become impossible to distinguish from translations. Similarly, positional imprecision makes it impossible to determine precise separate positions for exchange experiments; at Planck scales it is impossible to say whether particle exchange has taken place or not, and whether the wave function has changed sign or not. In short,

FIGURE 3 Andrei Sakharov (1921–1989).

▷ At Planck scales, spin cannot be defined or measured, and neither fermion nor boson behaviour can be defined or measured.

In particular, this implies that supersymmetry cannot be valid at Planck scales.

Challenge 37 e And we can continue. Due to measurement limitations, also *spatial parity* cannot be defined or measured at Planck scales.

We have thus shown that at Planck scales, particles do not interact locally, are not point-like, cannot be distinguished from antiparticles, cannot be distinguished from virtual particles, have no definite spin and have no definite spatial parity. We deduce that *particles do not exist at Planck scales.* Let us explore the remaining concept: particle mass.

A mass limit for elementary particles

The size d of any elementary particle must by definition be smaller than its own (reduced) Compton wavelength $\hbar/mc$. Moreover, the size of a particle is always larger than the Planck length: $d > l_{\mathrm{Pl}}$. Combining these two requirements and eliminating the size d, we get a constraint on the mass m of any elementary particle, namely

$$m < \frac{\hbar}{c\, l_{\mathrm{Pl}}} = \sqrt{\frac{\hbar c}{G}} = m_{\mathrm{Pl}} = 2.2 \cdot 10^{-8}\,\mathrm{kg} = 1.2 \cdot 10^{19}\,\mathrm{GeV}/c^2 \ . \tag{88}$$

The limit m_{Pl}, the so-called *Planck mass*, corresponds roughly to the mass of a human embryo that is ten days old, or equivalently, to that of a small flea. In short, *the mass of any elementary particle must be smaller than the Planck mass.* This fact was already Ref. 28 noted as 'well known' by Andrei Sakharov* in 1968; he explains that these hypothetical particles are sometimes called 'maximons'. And indeed, the known elementary particles all have masses well below the Planck mass. (In fact, the question why their masses are so very much smaller than the Planck mass is one of the most important questions of high-energy physics. We will come back to it.)

* Andrei Dmitrievich Sakharov, Soviet nuclear physicist (b. 1921 Moscow, d. 1989 Moscow). One of the keenest thinkers in physics, Sakharov, among others, invented the Tokamak, directed the construction of nuclear bombs, and explained the matter–antimatter asymmetry of nature. Like many others, he later campaigned against nuclear weapons, a cause for which he was put into jail and exile, together with his wife, Yelena Bonner. He received the Nobel Peace Prize in 1975.

There are many other ways to arrive at the mass limit for particles. For example, in order to measure mass by scattering – and that is the only way for very small objects – the Compton wavelength of the scatterer must be larger than the Schwarzschild radius; otherwise the probe will be swallowed. Inserting the definitions of the two quantities and neglecting the factor 2, we again get the limit $m < m_{\text{Pl}}$. In fact it is a general property of descriptions of nature that a minimum space-time interval leads to an upper limit for masses of elementary particles. Ref. 83

Farewell to massive particles – and to massless vacuum

The Planck mass divided by the Planck volume, i.e., the Planck density, is given by

$$\rho_{\text{Pl}} = \frac{c^5}{G^2\hbar} = 5.2 \cdot 10^{96}\,\text{kg/m}^3 \tag{89}$$

and is a useful concept in the following. One way to measure the (gravitational) mass M enclosed in a sphere of size R, and thus (roughly) of volume R^3, is to put a test particle in orbit around it at that same distance R. Universal gravitation then gives for the mass M the expression $M = Rv^2/G$, where v is the speed of the orbiting test particle. From $v < c$, we deduce that $M < c^2R/G$; since the minimum value for R is the Planck distance, we get (again neglecting factors of order unity) a limit for the mass density ρ, namely Challenge 38 e

$$\rho < \rho_{\text{Pl}} \,. \tag{90}$$

In other words, *the Planck density is the maximum possible value for mass density.*

Interesting things happen when we try to determine the error ΔM of a mass measurement in a Planck volume. Let us return to the mass measurement by an orbiting probe. From the relation $GM = rv^2$ we deduce by differentiation that $G\Delta M = v^2\Delta r + 2vr\Delta v > 2vr\Delta v = 2GM\Delta v/v$. For the error Δv in the velocity measurement we have the indeterminacy relation $\Delta v \geqslant \hbar/m\Delta r + \hbar/MR \geqslant \hbar/MR$. Inserting this in the previous inequality, and again forgetting the factor of 2, we find that the mass measurement error ΔM of a mass M enclosed in a volume of size R is subject to the condition

$$\Delta M \geqslant \frac{\hbar}{cR} \,. \tag{91}$$

Note that for everyday situations, this error is extremely small, and other errors, such as the technical limits of the balance, are much larger.

To check this result, we can explore another situation. We even use relativistic expressions, in order to show that the result does not depend on the details of the situation or the approximations. Imagine having a mass M in a box of size R, and weighing the box with a scale. (It is assumed that either the box is massless or that its mass is subtracted by the scale.) The mass error is given by $\Delta M = \Delta E/c^2$, where ΔE is due to the indeterminacy in the kinetic energy of the mass inside the box. Using the expression $E^2 = m^2c^4 + p^2c^2$, we get that $\Delta M \geqslant \Delta p/c$, which again reduces to equation (91). Now that we are sure of the result, let us continue.

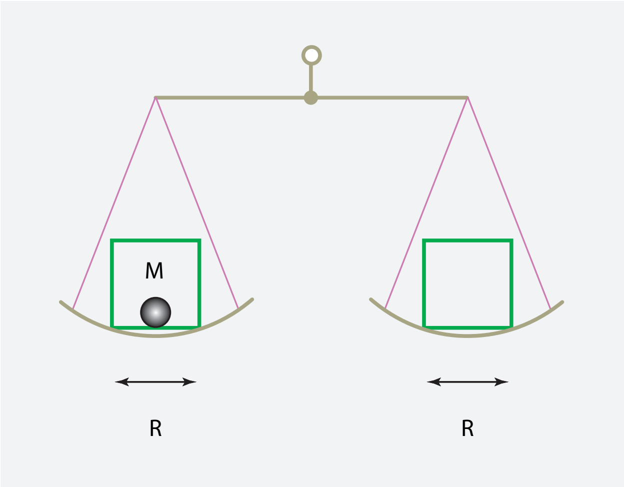


FIGURE 4 A thought experiment showing that matter and vacuum cannot be distinguished when the size of the enclosing box is of the order of a Planck length.

From equation (91) we deduce that for a box of Planck dimensions, the mass measurement *error* is given by the Planck mass. But from above we also know that the mass that can be put inside such a box must not be larger than the Planck mass. Therefore, for a box of Planck dimensions, the mass measurement *error* is larger than (or at best equal to) the mass contained in it: $\Delta M \geqslant M_{\rm Pl}$. In other words, if we build a balance with two boxes of Planck size, one empty and the other full, as shown in Figure 4, nature cannot decide which way the balance should hang! Note that even a repeated or a continuous measurement will not resolve the situation: the balance will change inclination at random, staying horizontal on average.

The argument can be rephrased as follows. The largest mass that we can put in a box of size R is a black hole with a Schwarzschild radius of the same value; the smallest mass present in such a box – corresponding to what we call a vacuum – is due to the indeterminacy relation and is given by the mass with a Compton wavelength that matches the size of the box. In other words, inside any box of size R we have a mass m, the limits of which are given by:

$$\underset{\text{(full box)}}{\frac{c^2 R}{G}} \geqslant m \geqslant \underset{\text{(empty box)}}{\frac{\hbar}{cR}} . \tag{92}$$

We see directly that for sizes R of the order of the Planck scales, the two limits coincide; in other words, we cannot distinguish a full box from an empty box in that case.

To be sure of this strange result, we check whether it also occurs if, instead of measuring the gravitational mass, as we have just done, we measure the inertial mass. The inertial mass for a small object is determined by touching it: physically speaking, by performing a scattering experiment. To determine the inertial mass inside a region of size R, a probe must have a wavelength smaller than R, and a correspondingly high energy. A high energy means that the probe also attracts the particle through gravity. (We thus find the intermediate result that *at Planck scales, inertial and gravitational mass can-*

not be distinguished. Even the balance experiment shown in Figure 4 illustrates this: at Planck scales, the two types of mass are always inextricably linked.) Now, in any scattering experiment, for example in a Compton-type experiment, the mass measurement is performed by measuring the wavelength change $\delta\lambda$ of the probe before and after the scattering. The mass indeterminacy is given by

$$\frac{\Delta M}{M} = \frac{\Delta\delta\lambda}{\delta\lambda} . \tag{93}$$

In order to determine the mass in a Planck volume, the probe has to have a wavelength of the Planck length. But we know from above that there is always a minimum wavelength indeterminacy, given by the Planck length l_{Pl}. In other words, for a Planck volume the wavelength error – and thus the mass error – is always as large as the Planck mass itself: $\Delta M \geqslant M_{\mathrm{Pl}}$. Again, this limit is a direct consequence of the limit on length and space measurements.

This result has an astonishing consequence. In these examples, the measurement error is independent of the mass of the scatterer: it is the same whether or not we start with a situation in which there is a particle in the original volume. We thus find that in a volume of Planck size, it is impossible to say whether or not there is something there when we probe it with a beam!

Matter and vacuum are indistinguishable

We can put these results in another way. On the one hand, if we measure the mass of a piece of vacuum of size R, the result is always at least $\hbar/cR$: there is no possible way to find a perfect vacuum in an experiment. On the other hand, if we measure the mass of a particle, we find that the result is size-dependent: at Planck scales it approaches the Planck mass for every type of particle, be it matter or radiation.

To use another image, when two particles approach each other to a separation of the order of the Planck length, the indeterminacy in the length measurements makes it impossible to say whether there is something or nothing between the two objects. In short, *matter and vacuum are interchangeable at Planck scales.* This is an important result: since mass and empty space cannot be differentiated, we have confirmed that they are made of the same 'fabric', of the same constituents. This idea, already suggested above, is now common to all attempts to find a unified description of nature.

This approach is corroborated by attempts to apply quantum mechanics in highly curved space-time, where a clear distinction between vacuum and particles is impossible, as shown by the Fulling–Davies–Unruh effect. Any accelerated observer, and any observer in a gravitational field, detects particles hitting him, even if he is in a vacuum. The effect shows that for curved space-time the idea of vacuum as particle-free space does not work. Since at Planck scales it is impossible to say whether or not space is flat, it is impossible to say whether it contains particles or not.

Ref. 84
Vol. V, page 129

In short, all arguments lead to the same conclusion: *vacuum, i.e., empty space-time, cannot be distinguished from matter at Planck scales.* Another common way to express this state of affairs is to say that when a particle of Planck energy travels through space it will be scattered by the fluctuations of space-time itself, as well as by matter, and the

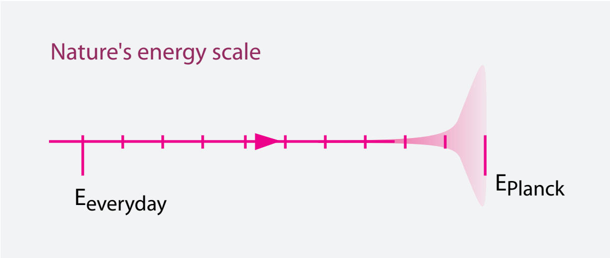


FIGURE 5 Planck effects make the energy axis an approximation.

two cases are indistinguishable. These surprising results rely on a simple fact: whatever definition of mass we use, it is always measured via combined length and time measurements. (This is even the case for normal weighing scales: mass is measured by the displacement of some part of the machine.) *Mass measurement is impossible at Planck scales.* The error in such mass measurements makes it *impossible to distinguish vacuum from matter.* In particular, *the concept of particle is not applicable at Planck scale.*

Curiosities and fun challenges on Planck scales

> There is nothing in the world but matter in motion, and matter in motion cannot move otherwise than in space and time.
> Lenin, *Materialism and empirio-criticism.*

Lenin's statement is wrong. And this is not so much because the world contains moving matter, moving radiation, moving vacuum and moving horizons, which is not exactly what Lenin claimed. Above all, his statement is wrong because at Planck scales, there is no matter, no radiation, no horizon, no space and no time. These concepts only appear at low energy. The rest of our adventure clarifies how.

* *

Observers are made of matter. Observers are not made of radiation. Observers are not made of vacuum. Observers are thus biased, because they take a specific standpoint. But at Planck scales, vacuum, radiation and matter cannot be distinguished. Two consequences follow: first, only at Planck scales would a description be free of any bias in favour of matter. Secondly, on the other hand, observers do not exist at all at Planck energy. Physics is thus only possible below Planck energy.

* *

If measurements become impossible near Planck energy, we cannot even draw a diagram with an energy axis reaching that value. A way out is shown Figure 5. The energy of elementary particles cannot reach the Planck energy.

* *

By the standards of particle physics, the Planck energy is rather large. Suppose we wanted to impart this amount of energy to protons using a particle accelerator. How large would a *Planck accelerator* have to be?
Challenge 39 s

* *

By the standards of everyday life, the Planck energy is rather small. Measured in litres of
Challenge 40 s gasoline, how much fuel does it correspond to?

* *

The usual concepts of matter and of radiation are not applicable at Planck scales. Usually, it is assumed that matter and radiation are made up of interacting elementary particles. The concept of an elementary particle implies an entity that is discrete, point-like, real and not virtual, has a definite mass and a definite spin, is distinct from its antiparticle, and, most of all, is distinct from vacuum, which is assumed to have zero mass. All these properties are lost at Planck scales. *At Planck scales, the concepts of 'mass', 'vacuum', 'elementary particle', 'radiation' and 'matter' do not make sense.*

* *

Do the large errors in mass measurements imply that mass can be negative at Planck
Challenge 41 s energy?

* *

We now have a new answer to the old question: why is there something rather than nothing? At Planck scales, there is *no difference* between something and nothing. We can now honestly say about ourselves that we are made of nothing.

* *

Special relativity implies that no length or energy can be invariant. Since we have come to the conclusion that the Planck energy and the Planck length are invariant, it appears that there must be deviations from Lorentz invariance at high energy. What effects would
Challenge 42 r follow? What kind of experiment could measure them? If you have a suggestion, pub-
Ref. 85 lish it! Several attempts are being explored. We will settle the issue later on, with some
Page 281 interesting insights.

* *

Quantum mechanics alone gives, via the Heisenberg indeterminacy relation, a lower limit to the *spread* of measurements, but, strangely enough, not on their *precision*, i.e.,
Ref. 86 not on the number of significant digits. Wolfgang Jauch gives an example: atomic lattice constants are known to a much higher precision than the positional indeterminacy of
Ref. 67 single atoms inside the crystal.

It is sometimes claimed that measurement indeterminacies smaller than the Planck values are possible for large enough numbers of particles. Can you show why this is
Challenge 43 s incorrect, at least for space and time?

* *

The idea that vacuum is not empty is not new. More than two thousand years ago, Aristotle argued for a filled vacuum, although his arguments were incorrect as seen from today's perspective. Also in the fourteenth century there was much discussion on whether empty space was composed of indivisible entities, but the debate died down again.

* *

A Planck-energy particle falling in a gravitational field would gain energy. But the Planck energy is the highest energy in nature. What does this apparent contradiction imply?
Challenge 44 s

* *

Ref. 59
One way to generalize the results presented here is to assume that, at Planck energy, nature is *event-symmetric*, i.e., symmetric under exchange of any two events. This idea, developed by Phil Gibbs, provides an additional formulation of the strange behaviour of nature at extreme scales.

* *

Vol. II, page 274
Because there is a minimum length in nature, so-called *singularities* do not exist. The issue, hotly debated for decades in the twentieth century, thus becomes uninteresting.

* *

Vol. V, page 146
Because mass and energy density are limited, any object of finite volume has only a finite number of degrees of freedom. The calculation of the entropy of black holes has confirmed that entropy values are always finite. This implies that perfect *baths* do not exist. Baths play an important role in thermodynamics (which must therefore be viewed as only an approximation), and also in recording and measuring devices: when a device measures, it switches from a neutral state to a state in which it shows the result of the measurement. In order not to return to the neutral state, the device must be coupled to a bath. Without a bath, a reliable measuring device cannot exist. In short, perfect clocks and length-measuring devices do not exist, because nature puts a limit on their storage ability.
Ref. 87

* *

Vol. I, page 26
If vacuum and matter cannot be distinguished, we cannot distinguish between objects and their environment. However, this was one of the starting points of our journey. Some interesting adventures still await us!

* *

Vol. III, page 323
We have seen earlier that characterizing nature as made up of particles and vacuum creates problems when interactions are included. On the one hand interactions are the difference between the parts and the whole, while on the other hand interactions are exchanges of quantum particles. This apparent contradiction can be used to show that something is counted twice in the usual characterization of nature. Noting that matter and space-time are both made of the same constituents resolves the issue.

* *

Challenge 45 d
Is there a smallest possible momentum? And a smallest momentum error?

* *

Given that time becomes an approximation at Planck scales, can we still ask whether nature is *deterministic*?

Let us go back to the basics. We can define time, because in nature change is not random, but gradual. What is the situation now that we know that time is only approximate? Is non-gradual change possible? Is energy conserved? In other words, are surprises possible in nature?

It is correct to say that time is not defined at Planck scales, and that therefore that determinism is an undefinable concept, but it is not a satisfying answer. What happens at 'everyday' scales? One answer is that at our everyday scales, the probability of surprises is so small that the world indeed is effectively deterministic. In other words, nature is not really deterministic, but the departure from determinism is not measurable, since every measurement and observation, by definition, *implies* a deterministic world. The lack of surprises would be due to the limitations of our human nature – more precisely, of our senses and brain.

Challenge 46 s Can you imagine any other possibility? In truth, it is not possible to prove these answers at this point, even though the rest of our adventure will do so. Page 427 We need to keep any possible alternative in mind, so that we remain able to check the answers.

* *

If matter and vacuum cannot be distinguished, then each has the properties of the other. For example, since space-time is an extended entity, matter and radiation are also extended entities. Furthermore, as space-time is an entity that reaches the borders of the system under scrutiny, particles must also do so. This is our first hint at the *extension of matter*; Page 117 we will examine this argument in more detail shortly.

* *

The impossibility of distinguishing matter and vacuum implies a lack of information at Planck scales. In turn, this implies an intrinsic basic entropy associated with any part of the universe at Planck scales. We will come back to this topic shortly, when we discuss Page 289 the entropy of black holes.

* *

Challenge 47 s When *can* matter and vacuum be distinguished? At what energy? This issue might be compared to the following question: Can we distinguish between a liquid and a gas by looking at a single atom? No, only by looking at many. Similarly, we cannot distinguish between matter and vacuum by looking at one point, but only by looking at many. We must always *average*. However, even averaging is not completely successful. Distinguishing matter from vacuum is like distinguishing clouds from the clear sky: like clouds, matter has no precise boundary.

* *

If the dimensionality of space is undefined at Planck scales, what does this mean for Challenge 48 e superstrings?

* *

Since vacuum, particles and fields are indistinguishable at Planck scales, we also lose the distinction between states and permanent, intrinsic properties of physical systems at those scales. This is a strong statement: the distinction was the starting point of our Vol. I, page 27 exploration of motion; the distinction allowed us to distinguish systems from their environment. In other words, *at Planck scales we cannot talk about motion!* This is a strong statement – but it is not unexpected. We are searching for the origin of motion, and we are prepared to encounter such difficulties.

Common constituents

> Es ist allerdings darauf hingewiesen worden, dass bereits die Einführung eines raum-zeitlichen Kontinuums angesischts der molekularen Struktur allen Geschehens im Kleinen möglicherweise als naturwidrig anzusehen sei. Vielleicht weise der Erfolg von Heisenbergs Methode auf eine rein algebraische Methode der Naturbeschreibung, auf die Ausschaltung kontinuierlicher Funktionen aus der Physik hin. Dann aber muss auch auf die Verwendung des Raum-Zeit-Kontinuums prinzipiell verzichtet werden. Es ist nicht undenkbar, dass der menschliche Scharfsinn einst Methoden finden wird, welche die Beschreitung dieses Weges möglich machen. Einstweilen aber erscheint dieses Projekt ähnlich dem Versuch, in einem luftleeren Raum zu atmen.*
>
> Albert Einstein, 1936, in *Physik und Realität*.

> One can give good reasons why reality cannot at all be represented by a continuous field. From the quantum phenomena it appears to follow with certainty that a finite system of finite energy can be completely described by a finite set of numbers (quantum numbers). This does not seem to be in accordance with a continuum theory, and must lead to an attempt to find a purely algebraic theory for the description of reality. But nobody knows how to obtain the basis of such a theory.
>
> Albert Einstein, 1955, the last sentences of *The Meaning of Relativity – Including the Relativistic Theory of the Non-Symmetric Field*, fifth edition. These were also his last published words.

In this rapid journey, we have destroyed all the experimental pillars of quantum theory: the superposition of wave functions, space-time symmetry, gauge symmetry, renormalization symmetry and permutation symmetry. We also have destroyed the foundations of special and general relativity, namely the concepts of the space-time manifold, fields, particles and mass. We have even seen that matter and vacuum cannot be distinguished.

It seems that we have lost every concept used for the description of motion, and thus made its description impossible. It seems that we have completely destroyed our two Page 70 'castles in the air', general relativity and quantum theory. And it seems that we are trying to breathe in airless space. Is this pessimistic view correct, or can we save the situation?

First of all, since matter and radiation are not distinguishable from vacuum, the quest for unification in the description of elementary particles is correct and necessary. There is no alternative to tearing down the castles and to continuing to breathe.

* 'Yet it has been suggested that the introduction of a space-time continuum, in view of the molecular structure of all events in the small, may possibly be considered as contrary to nature. Perhaps the success of Heisenberg's method may point to a purely algebraic method of description of nature, to the elimination of continuous functions from physics. Then, however, one must also give up, in principle, the use of the space-time continuum. It is not inconceivable that human ingenuity will some day find methods that will make it possible to proceed along this path. Meanwhile, however, this project resembles the attempt to breathe in an airless space.'

Page 70 See also what Einstein thought twenty years before. The new point is that he believes that an algebraic description is necessary. He repeats the point in the next quote.

Secondly, after tearing down the castles, the invariant Planck limits c, $\hbar$ and $c^4/4G$ still remain as a foundation.

Thirdly, after tearing down the castles, one important result appears. Since the concepts of 'mass', 'time' and 'space' cannot be distinguished from each other, a new, *single* entity or concept is necessary to define both particles and space-time. In short, vacuum and particles must be made of *common constituents*. In other words, we are not in airless space, and we uncovered the foundation that remains after we tore down the castles. Before we go on exploring these common constituents, we check what we have deduced so far against experiment.

Experimental predictions

A race is going on both in experimental and in theoretical physics: to be the first to suggest and to be the first to perform an experiment that detects a quantum gravity effect – apart possibly from (a part of) the Sokolov–Ternov effect. Here are some proposals.

Challenge 49 r
Vol. V, page 147

At Planck scales, space fluctuates. We might think that the fluctuations of space could blur the images of faraway galaxies, or destroy the phase relation between the photons. However, no blurring is observed, and the first tests show that light from extremely distant galaxies still interferes. The precise prediction of the phase washing effect is still being worked out; whatever the exact outcome, the effect is too small to be measured.

Ref. 88

Another idea is to measure the speed of light at different frequencies from faraway light flashes. There are natural flashes, called gamma-ray bursts, which have an extremely broad spectrum, from 100 GeV down to visible light at about 1 eV. These flashes often originate at cosmological distances d. Using *short* gamma-ray bursts, it is thus possible to test precisely whether the quantum nature of space-time influences the dispersion of light signals when they travel across the universe. Planck-scale quantum gravity effects *might* produce a dispersion. Detecting a dispersion would confirm that Lorentz symmetry breaks down at Planck scales.

Ref. 92, Ref. 91

The difference in arrival time Δt between two photon energies E_1 and E_2 defines a characteristic energy by

$$E_{\text{char}} = \frac{(E_1 - E_2)\, d}{c\, \Delta t} \,. \tag{94}$$

This energy value is between $1.4 \cdot 10^{19}$ GeV and over 10^{22} GeV for the best measurement to date. This is between just above the Planck energy and over one thousand times the Planck energy. However, despite this high characteristic energy, *no dispersion* has been found: even after a trip of ten thousand million years, all light arrives within one or two seconds.

Ref. 89, Ref. 90
Ref. 90

Another candidate experiment is the direct detection of distance fluctuations between bodies. Gravitational wave detectors are sensitive to extremely small noise signals in length measurements. There should be a noise signal due to the distance fluctuations induced near Planck energy. The indeterminacy in measurement of a length l is predicted to be

Ref. 92, Ref. 93
Ref. 94

$$\frac{\delta l}{l} \geqslant \left(\frac{l_{\text{Pl}}}{l}\right)^{2/3} \,. \tag{95}$$

This expression is deduced simply by combining the measurement limit of a ruler, from Page 67 quantum theory, with the requirement that the ruler not be a black hole. The sensitivity of the detectors to noise might reach the required level in the twenty-first century. The Ref. 95 noise induced by quantum gravity effects has also been predicted to lead to detectable quantum decoherence and vacuum fluctuations. So far, no such effect has been found.

A further candidate experiment for measuring quantum gravity effects is the detection of the loss of CPT symmetry at high energies. Especially in the case of the decay of Ref. 92 certain elementary particles, such as neutral kaons, the precision of experimental measurement is approaching the detection of Planck-scale effects. However, no such effect has been found yet.

Another possibility is that quantum gravity effects may change the threshold energy at Ref. 96 which certain particle reactions become possible. It may be that extremely high-energy photons or cosmic rays will make it possible to prove that Lorentz invariance is indeed broken at Planck scales. However, no such effect has been found yet.

In the domain of atomic physics, it has also been predicted that quantum gravity effects will induce a gravitational Stark effect and a gravitational Lamb shift in atomic Ref. 95 transitions. However, no such effect has been found yet.

Other proposals start from the recognition that the bound on the measurability of observables also puts a bound on the measurement *precision* for each observable. This bound is of no importance in everyday life, but it is important at Planck energy. One proposal is to search for a minimal noise in length measurements, e.g., in gravitational wave detectors. But no such noise has been found yet.

In summary, the experimental detection of quantum gravity effects *might* be possible, despite their weakness, at some time during the twenty-first century. The successful prediction and detection of such an effect would be one of the highlights of physics, as it would challenge the usual description of space and time even more than general relativ- Ref. 97 ity did. On the other hand, most unified models of physics predict the *absence* of any measurable quantum gravity effect.

SUMMARY ON PARTICLES AND VACUUM

Combining quantum theory and general relativity leads us to several important results on the description of nature:

— *Vacuum and particles and horizons mix* at Planck scales, because there is no conceivable way to distinguish whether a Planck-sized region is part of a particle or of empty space or of a horizon. Matter, radiation and vacuum cannot be distinguished at Planck scales. Equivalently, empty space and particles and horizons are made of Ref. 98 *fluctuating common constituents*.

— We note that all arguments of this chapter equally imply that vacuum and particles and horizons mix *near* Planck scales. For example, matter, radiation and vacuum cannot be distinguished *near* Planck scales.

— The constituents of vacuum and particles *cannot be points*. There is no conceivable way to prove that points exist, because the smallest measurable distance in nature is the Planck length.

— Particles, vacuum and continuous space *do not exist* at Planck scales. They disappear in a yet unclear Planck scale mixture.
— The three independent Planck limits c, $\hbar$ and $c^4/4G$ *remain valid* also in domains where quantum theory and general relativity are combined.

All these results must be part of the complete theory that we are looking for. Generally speaking, we found the same conclusions that we found already in the chapter on limit statements. We thus continue along the same path that we took back then: we explore the universe as a whole.

Page 55

Chapter 5

WHAT IS THE DIFFERENCE BETWEEN THE UNIVERSE AND NOTHING?

> Die Grenze ist der eigentlich fruchtbare Ort der Erkenntnis.**
>
> Paul Tillich, *Auf der Grenze.*

This strange question is the topic of the current leg of our mountain ascent. In the last section we explored nature in the vicinity of Planck scales. In fact, the other limit, namely the description of motion at large, cosmological scales, is equally fascinating. As we proceed, many incredible results will appear, and at the end we will discover the answer to the question in the title.

Cosmological scales

> Hic sunt leones.***
>
> Antiquity

The description of motion requires the application of general relativity whenever the scale d of the situation is of the order of the Schwarzschild radius, i.e., whenever

$$d \approx r_S = 2Gm/c^2 . \quad (96)$$

It is straightforward to confirm that, with the usually quoted mass m and size d of everything visible in the universe, this condition is indeed fulfilled. We do need general relativity, and thus curved space-time, when talking about the whole of nature. Challenge 50 s

Similarly, quantum theory is required for the description of the motion of an object whenever we approach it within a distance d of the order of the (reduced) Compton wavelength $\lambda\!\!\!^{-}_C$, i.e., whenever

$$d \approx \lambda\!\!\!^{-}_C = \frac{\hbar}{mc} . \quad (97)$$

Obviously, for the total mass of the universe this condition is not fulfilled. However, we are not interested in the motion of the universe itself; we are interested in the motion of its components. In the description of these components, quantum theory is required whenever pair production and annihilation play a role. This is the case in the early history

** 'The frontier is the really productive place of understanding.' Paul Tillich (b. 1886 Starzeddel, d. 1965 Chicago), theologian, socialist and philosopher.

*** 'Here are lions.' This was written across unknown and dangerous regions on ancient maps.

of the universe and near the cosmological horizon, i.e., for the most distant events that we can observe in space and time. We are thus obliged to include quantum theory in any precise description of the universe.

Since at cosmological scales we need both quantum theory and general relativity, we start our investigation with the study of time, space and mass, by asking at large scales the same questions that we asked above at Planck scales.

Maximum time

Is it possible to measure time intervals of any imaginable size? Observations show that in nature there is a maximum time interval t_0, with a value of about

$$t_0 = 435\,\mathrm{Ps} \text{ or } 13\,800 \text{ million years,}$$

and providing an upper limit to the measurement of time. It is called the *age of the universe*, and has been deduced from two sets of measurements: the expansion of space-time and the age of matter.

We are all familiar with clocks components that have been ticking for a long time: the hydrogen atoms in our body. All hydrogen atoms were formed just after the big bang. We can almost say that the electrons in these atoms have been orbiting their nuclei since the dawn of time. In fact, the quarks inside the protons in these atoms have been moving a few hundred thousand years longer than the electrons.

We thus have an upper time limit for any clock made of atoms. Even 'clocks' made of radiation (can you describe one?) yield a similar maximum time. Now, the study of the spatial expansion of the universe leads to the *same* maximum age. No clock or measurement device was ticking *longer ago* than this maximum time, and no clock could provide a record of having done so.

Challenge 51 s

In summary, it is not possible to measure time intervals greater than the maximum time, either by using the history of space-time or by using the history of matter or radiation.* The maximum time is thus rightly called the *age* of the universe. Of course, this is not new. But exploring the age issue in more detail does reveal some surprises.

Does the universe have a definite age?

> One should never trust a woman who tells one her real age. A woman who would tell one that, would tell one anything.
>
> Oscar Wilde**

In light of all measurements, it may seem silly to question the age of the universe. The age value is found in many books and tables and its precise determination is one of the most important quests in modern astrophysics. But is this quest reasonable?

Vol. II, page 307

In order to measure the duration of a movement or the age of a system, we need a clock that is *independent* of that movement or system, and thus *outside* the system. How-

* This implies that so-called 'oscillating universe' models, in which it is claimed that 'before' the big bang there were other phenomena, cannot be justified on the basis of nature or observations. They are based on beliefs.

** Oscar Wilde, (b. 1854 Dublin, d. 1900 Paris), poet and playwright, equally famous for his wit.

ever, there are no clocks outside the universe, and no clock inside it can be independent. In fact, we have just seen that no clock inside the universe can run throughout its full history. In particular, no clock can run through its earliest history.

Time can be defined only if it is possible to distinguish between matter and space. Given this distinction, we can talk either about the age of *space*, by assuming that matter provides suitable and independent clocks – as is done in general relativity – or about the age of *matter*, such as stars or galaxies, by assuming that the expansion of space-time, or possibly of other matter, provides a good clock. Both possibilities are being explored experimentally in modern astrophysics – and both give the same result, of about fourteen thousand million years, which was mentioned above. Despite this correspondence, for the universe as a *whole*, an age *cannot* be defined, because there is no clock outside it!

The issue of the starting point of time makes this difficulty even more apparent. We may imagine that going back in time leads to only two possibilities: either the starting instant $t = 0$ is part of time or it is not. (Mathematically, this means that the segment representing time is either closed or open.) Both these possibilities imply that it is possible to measure arbitrarily small times; but we know from the combination of general relativity and quantum theory that this is *not* the case. In other words, neither possibility is correct: the beginning cannot *be* part of time, nor can it *not be* part of it. There is only one solution to this contradiction: *there was no beginning at all.*

The lack of a beginning of time is consistent with a minimum length and a minimum action. Indeed, both limits imply that there is a maximum curvature for space-time. Curvature can be measured in several ways: for example, surface curvature is an inverse area. Within a factor of order one, we find

$$K < \frac{c^3}{G\hbar} = 0.39 \cdot 10^{70}\ \mathrm{m}^{-2} \tag{98}$$

as a limit for the surface curvature K in nature. In other words, the universe has never been as small as a point, never had zero age, never had infinite density, and never had infinite curvature. It is not difficult to get a similar limit for temperature or any other physical quantity near the big bang. Challenge 52 s

There is a more drastic formulation. Since gravity and quantum effect imply that events do not exist in nature,

▷ The big bang cannot have been an event.

There never was an initial singularity or a beginning of the universe.

In short, the situation is consistently muddled. Neither the age of the universe nor its origin makes sense. What is going wrong? Or rather, *how* are things going wrong? What happens if instead of jumping directly to the big bang, we *approach* it as closely as possible? To clarify the issue, we ask about the measurement *error* in our statement that the universe is fourteen thousand million years old. This turns out to be a fascinating topic.

How precise can age measurements be?

> No woman should ever be quite accurate about her age. It looks so calculating.
>
> Oscar Wilde

The first way to measure the age of the universe* is to look at clocks in the usual sense of the word, namely at clocks made of *matter*. As explained in the part on quantum theory, Salecker and Wigner showed that a clock built to measure a total time T with a precision Δt has a minimum mass m given by

Vol. V, page 46
Ref. 55, Ref. 56

$$m > \frac{\hbar}{c^2} \frac{T}{(\Delta t)^2} . \tag{99}$$

A simple way to incorporate general relativity into this result was suggested by Ng and van Dam. Any clock of mass m has a minimum resolution Δt due to the curvature of space that it introduces, given by

Ref. 94

$$\Delta t > \frac{Gm}{c^3} . \tag{100}$$

If m is eliminated, these two results imply that a clock with a precision Δt can only measure times T up to a certain maximum value, namely

$$T < \frac{(\Delta t)^3}{t_{\text{Pl}}^2} , \tag{101}$$

where $t_{\text{Pl}} = \sqrt{\hbar G/c^5} = 5.4 \cdot 10^{-44}$ s is the Planck time. (As usual, we have omitted factors of order one in this and in all the following results of this chapter.) In other words, the higher the accuracy of a clock, the shorter the time during which it works dependably. The precision of a clock is limited not only by the expense of building it, but also by nature itself. Nevertheless, it is easy to check that for clocks used in daily life, this limit is not even remotely approached. For example, you may wish to calculate how precisely your own age can be specified.

Challenge 53 e

As a consequence of the inequality (101), a clock trying to achieve an accuracy of one Planck time can do so for at most one Planck time! *A real clock cannot achieve Planck-time accuracy.* If we try to go beyond the limit (101), fluctuations of space-time hinder the working of the clock and prevent higher precision. With every Planck time that passes, the clock accumulates a measurement error of at least one Planck time. Thus, *the total measurement error is at least as large as the measurement itself.* This conclusion is also valid for clocks based on radiation.

In short, measuring age with a clock always involves errors. Whenever we try to reduce these errors to the smallest possible level, the Planck level, the clock becomes so imprecise over large times that age measurements become impossible.

* The age t_0 is not the same as the Hubble time $T = 1/H_0$. The Hubble time is only a computed quantity and (almost) always larger than the age; the relation between the two depends on the values of the cosmological constant, the density and other properties of the universe. For example, for the standard 'hot big bang' scenario, i.e., for the matter-dominated Einstein–de Sitter model, we have the simple relation $T = (3/2)\, t_0$.

Ref. 99

Does time exist?

> Time is waste of money.
>
> Oscar Wilde

Vol. I, page 40

Ever since people began to study physics, the concept of 'time' has designated what is measured by a clock. But the inequality (101) for a maximum clock time implies that perfect clocks do not exist, and thus that time is only an approximate concept: perfect time does not exist. Thus, in nature there is no 'idea' of time, in the Platonic sense. In fact, the discussion so far can be seen as proof that combining quantum theory and general relativity, because of the resulting measurement errors, prevents the existence of perfect or 'ideal' examples of any classical observable or any everyday concept.

Challenge 54 e

Time does not exist. Yet it is obviously a useful concept in everyday life. The key to understanding this is *measurement energy*. Any clock – in fact, any system of nature – is characterized by a simple number, namely the highest ratio of its kinetic energy to the rest energy of its components. In daily life, this ratio is about $1\,\mathrm{eV}/10\,\mathrm{GeV} = 10^{-10}$. Such *low-energy* systems are well suited for building clocks. The more precisely the motion of the main moving part – the pointer of the clock – can be kept constant and monitored, the higher the precision of the clock. To achieve very high precision, the pointer must have very high mass. Indeed, in any clock, both the position and the speed of the pointer must be measured, and the two measurement errors are related by the quantum-mechanical indeterminacy relation $\Delta v\,\Delta x > \hbar/m$. High mass implies low intrinsic fluctuation. Furthermore, in order to screen the pointer from outside influences, even more mass is needed. This connection between mass and accuracy explains why more accurate clocks are usually more expensive.

Page 59 Page 69 Challenge 55 e

The standard indeterminacy relation $m\Delta v\,\Delta x > \hbar$ is valid only at everyday energies. However, we cannot achieve ever higher precision simply by increasing the mass without limit, because general relativity changes the indeterminacy relation to $\Delta v\,\Delta x > \hbar/m + G(\Delta v)^2 m/c^3$. The additional term on the right-hand side, negligible at everyday scales, is proportional to energy. Increasing it by a large amount limits the achievable precision of the clock. The smallest measurable time interval turns out to be the Planck time.

In summary,

▷ Time exists, as a good approximation, only for *low-energy* systems.

Any increase in precision beyond a certain limit requires an increase in the energy of the components; at Planck energy, this increase will prevent an increase in precision.

What is the error in the measurement of the age of the universe?

Challenge 56 e

It is now straightforward to apply our discussion about the measurement of time to the age of the universe. The inequality (101) implies that the highest precision possible for a clock is about 10^{-23} s, or about the time light takes to move across a proton. The finite age of the universe also yields a maximum *relative* measurement precision. Inequality (101) can be written as

$$\frac{\Delta t}{T} > \left(\frac{t_{\mathrm{Pl}}}{T}\right)^{2/3} . \qquad (102)$$

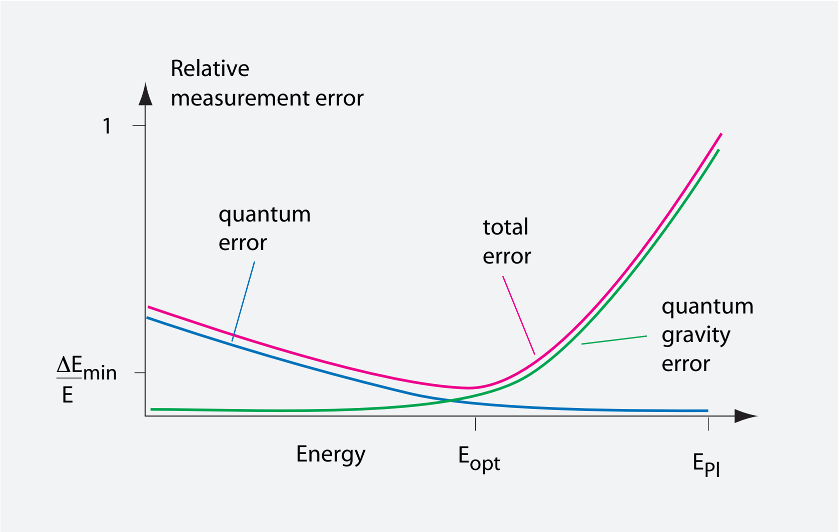


FIGURE 6 Measurement errors as a function of measurement energy.

Inserting the age of the universe for T, we find that no time interval can be measured with a precision of more than about 40 decimals.

To clarify the issue, we can calculate the error in measurement as a function of the observation energy E_{meas}, the energy of the measurement probe. There are two limit cases. For *low* energies, the error is due to quantum effects and is given by

$$\frac{\Delta t}{T} \sim \frac{1}{E_{meas}} \tag{103}$$

which decreases with increasing measurement energy. For *high* energies, however, the error is due to gravitational effects and is given by

$$\frac{\Delta t}{T} \sim \frac{E_{meas}}{E_{Pl}} \tag{104}$$

so that the total error varies as shown in Figure 6. In particular, very high energies do not reduce measurement errors: any attempt to reduce the measurement error for the age of the universe below 10^{-23} s would require energies so high that the limits of space-time would be reached, making the measurement itself impossible. We reached this conclusion through an argument based on clocks made of particles. We will see below that trying to determine the age of the universe from its expansion leads to the same limitation.

Imagine observing a tree which, as a result of some storm or strong wind, has fallen towards second tree, touching it at the very top, as shown in Figure 7. It is possible to determine the heights of both trees by measuring their separation and the angles at the base. The *error* in the heights will depend on the errors in measurement of the separation and angles.

Similarly, the age of the universe can be calculated from the present distance and

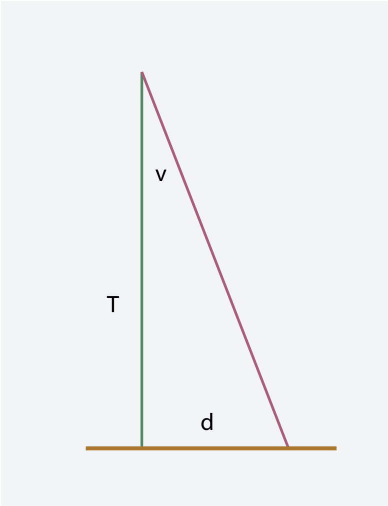


FIGURE 7 Trees and galaxies.

speed of objects – such as galaxies – observed in the night sky. The present distance d corresponds to separation of the trees at ground level, and the speed v to the angle between the two trees. The Hubble time T of the universe (which is usually assumed to be larger than the age of the universe) then corresponds to the height at which the two trees meet. This age – in a naive sense, the time since the galaxies 'separated' – is given, within a factor of order one, by

$$T = \frac{d}{v} . \tag{105}$$

In simple terms, this is the method used to determine the age of the universe from the expansion of space-time, for galaxies with red-shifts below unity.* The (positive) measurement error ΔT becomes

$$\frac{\Delta T}{T} = \frac{\Delta d}{d} + \frac{\Delta v}{v} . \tag{106}$$

Let us explore this in more detail. For any measurement of T, we have to choose the object, i.e., a distance d, as well as an observation time Δt, or, equivalently, an observation energy $\Delta E = 2\pi\hbar/\Delta t$. We will now investigate the consequences of these choices for equation (106), always taking into account both quantum theory and general relativity.

At everyday energies, the result of the determination of the age of the universe t_0 is about $(13.8 \pm 0.1) \cdot 10^9$ Ga. This value is deduced by measuring red-shifts, i.e., velocities, and distances, using stars and galaxies in distance ranges, from some hundred thousand light years up to a red-shift of about 1. Measuring red-shifts does not produce large velocity errors. The main source of experimental error is the difficulty in determining the distances of galaxies.

* At higher red-shifts, the speed of light, as well as the details of the expansion, come into play. To continue with the analogy of the trees, we find that the trees are not straight all the way up to the top and that they grow on a slope, as suggested by Figure 8.

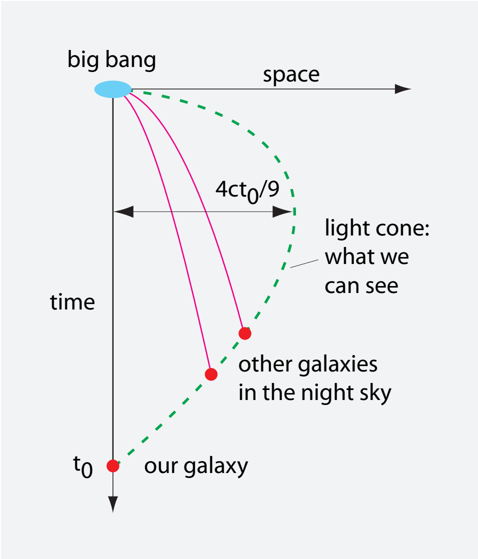


FIGURE 8 The speed and distance of remote galaxies.

What is the smallest possible error in distance? Obviously, inequality (102) implies

$$\frac{\Delta d}{T} > \left(\frac{l_{\rm Pl}}{d}\right)^{2/3} \tag{107}$$

thus giving the same indeterminacy in the age of the universe as the one we found above in the case of material clocks. Challenge 57 e

We can try to reduce the age error in two ways: by choosing objects at either small or large distances. Let us start with small distances. In order to get high precision at small distances, we need high observation energies. It is fairly obvious that at observation energies near the Planck value, $\Delta T/T$ approaches unity. In fact, both terms on the right-hand side of equation (106) become of order one. At these energies, Δv approaches c and the maximum value for d approaches the Planck length, for the same reason that at Planck energy the maximum measurable time is the Planck time. In short, Challenge 58 e

▷ At Planck scales it is impossible to say whether the universe is old or young.

Let us consider the other extreme, namely objects extremely far away, say with a red-shift of $z \gg 1$. Relativistic cosmology requires the diagram of Figure 7 to be replaced by the more realistic diagram of Figure 8. The 'light onion' replaces the familiar light cone of special relativity: light converges near the big bang. In this case the measurement error for the age of the universe also depends on the distance and velocity errors. At the largest possible distances, the signals an object sends out must be of high energy, because the emitted wavelength must be smaller than the universe itself. Thus, inevitably, we reach Planck energy. However, we have seen that in such high-energy situations, both the emitted radiation and the object itself are indistinguishable from the space-time background. In other words, the red-shifted signal we would observe today would have Ref. 99

a wavelength as large as the size of the universe, with a correspondingly small frequency.

There is another way to describe the situation. At Planck energy or near the horizon, the original signal has an error of the same size as the signal itself. When measured at the present time, the red-shifted signal still has an error of the same size as the signal. As a result, the error in the horizon distance becomes as large as the value to be measured.

In short, even if space-time expansion and large scales are used, the instant of the so-called beginning of the universe cannot be determined with an error smaller than the age of the universe itself: a result we also found at Planck distances. If we aim for perfect precision, we just find that the universe is 13.8 ± 13.8 thousand million years old! In other words, *in both extremal situations, it is impossible to say whether the universe has a non-vanishing age*.

We have to conclude that the anthropocentric concept of 'age' does not make any sense for the universe as a whole. The usual textbook value is useful only for ranges of time, space and energy in which matter and space-time are clearly distinguished, namely at everyday, human-scale energies; the value has no more general meaning.

Challenge 59 ny You may like to examine the issue of the *fate* of the universe using the same arguments. But we will now continue on the path outlined at the start of this chapter; the next topic on this path is the measurement of length.

Maximum length

General relativity shows that the *horizon distance*, i.e., the distance of objects with infinite red-shift, is finite. In the usual cosmological model, for hyperbolic (open) and parabolic (marginal) evolutions of the universe, the *size* of the universe is assumed infinite.* For elliptical evolution, the total size is finite and depends on the curvature. However, in this Vol. II, page 243 case also the present measurement limit yields a minimum size for the universe many times larger than the horizon distance.

Quantum field theory, on the other hand, is based on flat and infinite space-time. Let us see what happens when the two theories are combined. What can we say about measurements of length in this case? For example, would it be possible to construct and use a metre rule to measure lengths larger than the distance to the horizon?

Admittedly, we would have no time to push the metre rule out up to the horizon, Ref. 99 because in the standard big bang model the horizon moves away from us faster than the speed of light. (We should have started using the metre rule right at the big bang.) But just for fun, let us assume that we have actually managed to do this. How far away can we read off distances? In fact, since the universe was smaller in the past, and since every observation of the sky is an observation of the past, Figure 8 shows that the maximum Ref. 99 *spatial distance* away from us at which an object can be seen is only $4ct_0/9$. Obviously, for space-time intervals, the maximum remains ct_0.

Thus, in all cases it turns out to be impossible to measure lengths larger than the horizon distance, even though general relativity sometimes predicts such larger distances.

* In cosmology, we need to distinguish between the scale factor R, the Hubble radius $c/H = cR/\dot{R}$, the horizon distance h and the size d of the universe. The Hubble radius is a computed quantity giving the distance at which objects move away with the speed of light. The Hubble radius is always *smaller* than the horizon distance, at which in the standard Einstein–de Sitter model, for example, objects move away with Ref. 99 *twice* the speed of light. However, the horizon itself moves away with *three* times the speed of light.

This result is unsurprising, and in obvious agreement with the existence of a limit for measurements of time intervals. The real surprises come next.

Is the universe really a big place?

Ref. 100 Astronomers and Hollywood films answer this question in the affirmative. Indeed, the Vol. II, page 307 distance to the horizon of the universe is often included in tables. Cosmological models specify that the scale factor R, which fixes the distance to the horizon, grows with time t; for the case of the standard mass-dominated Einstein–de Sitter model, i.e., for a vanishing cosmological constant and flat space, we have

$$R(t) = C\, t^{2/3} , \tag{108}$$

where the numerical constant C relates the commonly accepted horizon distance to the commonly accepted age. Indeed, observation shows that the universe is large, and is getting larger. But let us investigate what happens if we add some quantum theory to this result from general relativity. Is it really possible to measure the distance to the horizon?

We look first at the situation at high (probe) energies. We saw above that space-time and matter are not distinguishable at Planck scales. Therefore, at Planck energy we cannot state whether or not objects are *localized*. At Planck scales, the distinction between matter and vacuum – so basic to our thinking – disappears.

Another way to say this is that we cannot claim that space-time is *extended* at Planck scales. Our concept of extension derives from the possibility of measuring distances and time intervals, and from observations such as the ability to align several objects behind one another. Such observations are not possible at Planck scales and energies, because of the inability of probes to yield useful results. In fact, all of the everyday observations from which we deduce that space is extended are impossible at Planck scales, where *the basic distinction between vacuum and matter, namely between extension and localization, disappears*. As a consequence, at Planck energy the size of the universe cannot be measured. It cannot even be called larger than a matchbox.

The problems encountered with probes of high probe energies have drastic consequences for the size measurement of the universe. All the arguments given above for Challenge 60 e the errors in measurement of the age can be repeated for the distance to the horizon. To reduce size measurement errors, a measurement probe needs to have high energy. But at high energy, measurement errors approach the value of the measurement results. At the largest distances and at Planck energy, the measurement errors are of the same magnitude as the measured values. If we try to determine the size of the universe with high precision, we get no precision at all.

The inability to get precise values for the size of the universe should not come unexpected. For a reliable measurement, the standard must be different, independent, and outside the system to be measured. For the universe this is impossible.

Studying the size of the big bang also produces strange results. The universe is said to have been much smaller near the big bang because, on average, all matter is moving away from all other matter. But if we try to follow the path of matter into the past with high precision, using Planck energy probes, we get into trouble: since measurement errors are as large as measurement data, we cannot claim that the universe was smaller near the big bang than it is today: there is no way to reliably distinguish size values.

There are other confirmations too. If we had a metre rule spanning the whole universe, even beyond the horizon, with zero at the place where we live, what measurement *error* would it produce for the horizon? It does not take long to work out that the expansion of space-time, from Planck scales to the present size, implies an expansion in the error from Planck size to a length of the order of the present distance to the horizon. Again, the error is as large as the measurement result. And again, the size of the universe turns out not to be a meaningful property.

Challenge 61 e

Since this reasoning also applies if we try to measure the diameter of the universe instead of its radius, it is impossible to say whether the antipodes in the sky really are distant from each other!

We can summarize the situation by noting that anything said about the size of the universe is as limited as anything said about its age. *The height of the sky depends on the observation energy.* If we start measuring the sky at standard observation energies, and try to increase the precision of measurement of the distance to the horizon, the measurement error increases beyond all bounds. At Planck energy, the volume of the universe is indistinguishable from the Planck volume – and vice versa.

The boundary of space – is the sky a surface?

The horizon of the universe – essentially, the black part of the night sky – is a fascinating entity. Everybody interested in cosmology wants to know what happens there. In newspapers the horizon is sometimes called the *boundary of space*. Some surprising insights – which have not yet made it to the newspapers – appear when we combine general relativity and quantum mechanics.

We have seen that the errors in measuring the distance of the horizon are substantial. They imply that we cannot pretend that all points of the sky are equally far away from us. Thus we cannot say that the sky is a surface; it could be a volume. In fact, there is no way to determine the dimensionality of the horizon, nor the dimensionality of space-time near the horizon.*

Thus measurements do not allow us to determine whether the boundary is a point, a surface, or a line. It may be a very complex shape, even knotted. In fact, quantum theory tells us that it must be all of these from time to time: that *the sky fluctuates in height and shape*.

In short, measurement errors prevent the determination of the topology of the sky. In fact, this is not a new result. As is well known, general relativity is unable to describe particle–antiparticle pair creation particles with spin 1/2. The reason for this inability is the change in space-time topology required by such processes. The universe is full of these and many other quantum processes; they imply that it is impossible to determine or define the microscopic topology for the universe and, in particular, for the horizon. Can you find at least two other arguments to confirm this conclusion?

Challenge 63 s

* The measurement errors also imply that we cannot say anything about translational symmetry at cosmological scales. Can you confirm this? In addition, at the horizon it is impossible to distinguish between space-like and time-like distances. Even worse, concepts such as 'mass' or 'momentum' become muddled at the horizon. This means that, as at Planck energy, we are unable to distinguish between object and background, and between state and intrinsic properties. We thus confirm the point made above.

Challenge 62 ny
Page 86

Worse still, quantum theory shows that space-time is not continuous at a horizon: this can easily be deduced using the Planck-scale arguments from the previous section. Page 59 Time and space are not defined at horizons.

Finally, there is no way to decide by measurement whether the various points on the horizon are *different* from each other. On the horizon, measurement errors are of the same order as the size of the horizon. The distance between two points in the night sky is thus undefined. Therefore it is unclear what the *diameter* of the horizon is.

In summary, the horizon has no specific distance or shape. The horizon, and thus the universe, cannot be shown to be manifolds. This unexpected result leads us to a further question.

Does the universe have initial conditions?

Ref. 101 One often reads about the quest for the initial conditions of the universe. But before joining this search, we should ask *whether* and *when* such initial conditions make any sense.

Obviously, our everyday description of motion requires knowledge of initial conditions, which describe the *state* of a system, i.e., all those aspects that differentiate it from a system with the same intrinsic properties. Initial conditions – like the state – are attributed to a system by an *outside* observer.

Quantum theory tells us that initial conditions, or the state of a system, can only be defined by an outside observer with respect to an environment. It is already difficult to be outside the universe – but even inside the universe, a state can only be defined if matter can be distinguished from vacuum. This is impossible at Planck energy, near the big bang, or at the horizon. Thus

▷ The universe has no state. And thus *no wave function*.

Page 59 The limits imposed by the Planck values confirm this conclusion in other ways. First of all, they show that the big bang was not a singularity with infinite curvature, density or temperature, because infinitely large values do not exist in nature. Secondly, since instants of time do not exist, it is impossible to define the state of any system at a given time. Thirdly, as instants of time do not exist, neither do events, and so the big bang was not an event, and neither an initial state nor an initial wave function can be ascribed to the universe. (Note that this also means that the universe cannot have been created.)

Page 308 In short, *there are no initial conditions for the universe*. Initial conditions make sense only for subsystems, and only far from Planck scales. Thus, for initial conditions to exist, the system must be far from the horizon and it must have evolved for some time 'after' the big bang. Only when these two requirements are fulfilled can objects *move* in space. Of course, this is always the case in everyday life. The lack of initial conditions means Page 19 that we have (partly) solved one issue from the millennium list.

At this point in our mountain ascent, where neither time nor length is clearly defined at cosmological scales, it should come as no surprise that there are similar difficulties concerning the concept of mass.

Does the universe contain particles and stars?

The total number of stars in the universe, about $10^{23\pm1}$, is included in every book on cosmology. A smaller number can be counted on clear nights. But how dependable is the statement? We can ask the same question about particles instead of stars. The commonly quoted numbers are $10^{80\pm1}$ baryons and $10^{89\pm1}$ photons. However, the issue is not simple. Neither quantum theory nor general relativity alone make predictions about the number of particles. What happens if we combine the two theories?

Vol. II, page 307

In order to define the number of particles in a region, quantum theory first of all requires a vacuum state to be defined. The number of particles is defined by comparing the system with the vacuum. If we neglect or omit general relativity by assuming flat space-time, this procedure poses no problem. However, if we include general relativity, and thus a curved space-time, especially one with a strangely behaved horizon, the answer is simple: there is *no* vacuum state with which we can compare the universe, for two reasons. First, nobody can explain what an empty universe would look like. Secondly, and more importantly, there is no way to define a state of the universe. The number of particles in the universe thus becomes undefinable. Only at everyday energies and for finite dimensions are we able to speak of an approximate number of particles.

Vol. IV, page 115

In the case of the universe, a comparison with the vacuum is also impossible for purely practical reasons. The particle counter would have to be outside the system. (Can you confirm this?) In addition, it is impossible to remove particles from the universe.

Challenge 64 e

The impossibility of defining a vacuum state, and thus the number of particles in the universe, is not surprising. It is an interesting exercise to investigate the measurement errors that appear when we try to determine the number of particles despite this fundamental impossibility.

Challenge 65 s

Can we count the stars? In principle, the same conclusion applies as for particles. However, at everyday energies the stars can be counted *classically*, i.e., without taking them out of the volume in which they are enclosed. For example, this is possible if the stars are differentiated by mass, colour or any other individual property. Only near Planck energy or near the horizon are these methods inapplicable. In short, the number of stars is only defined as long as the observation energy is low, i.e., as long as we stay away from Planck energy and from the horizon.

So, despite appearances on human scales,

▷ There is no definite number of particles in the universe.

The universe cannot be distinguished from vacuum by counting particles. Even though particles are necessary for our own existence and functioning, a complete count of all of them cannot be made.

This conclusion is so strange that we should try to resist it. Let us try another method of determining the content of matter in the universe: instead of counting particles in the universe, let us weigh the universe.

Does the universe have mass?

Mass distinguishes objects from the vacuum. The average mass density of the universe, about 10^{-26} kg/m^3, is often cited in texts. Is it different from a vacuum? Quantum theory

Vol. II, page 307

shows that, as a result of the indeterminacy relation, even an empty volume of size R has a mass. For a zero-energy photon inside such a vacuum, we have $E/c = \Delta p > \hbar/\Delta x$, so that in a volume of size R, we have a minimum mass of at least $m_{\min}(R) = h/cR$. For a spherical volume of radius R there is thus a minimal mass density given approximately by

$$\rho_{\min} \approx \frac{m_{\min}(R)}{R^3} = \frac{\hbar}{cR^4} . \tag{109}$$

For the universe, if the standard horizon distance R_0 of 13 800 million light years is inserted, the value becomes about $10^{-142}\,\mathrm{kg/m^3}$. This describes the density of the vacuum. In other words, the universe, with a textbook density of about $10^{-26}\,\mathrm{kg/m^3}$, seems to be clearly different from vacuum. But are we sure?

We have just deduced that the radius of the horizon is undefined: depending on the observation energy, it can be as small as the Planck length. This implies that the density of the universe lies somewhere between the lowest possible value, given by the density of vacuum $\rho_{\min}$ just mentioned, and the highest possible one, namely the Planck density.* In short, the total mass of the universe depends on the energy of the observer.

Vol. I, page 100
Vol. II, page 65
Another way to measure the mass of the universe would be to apply the original definition of mass, as given in classical physics and as modified by special relativity. Thus, let us try to collide a standard kilogram with the universe. It is not hard to see that whatever we do, using either low or high energies for the standard kilogram, the mass of the universe cannot be constrained by this method. Challenge 67 e We would need to produce or to measure a velocity change Δv for the rest of the universe after the collision. To hit all the mass in the universe at the same time, we need high energy; but then we are hindered by Planck energy effects. In addition, a properly performed collision measurement would require a mass outside the universe, which is rather difficult to achieve.

Yet another way to measure the mass would be to determine the gravitational mass of the universe through straightforward weighing. But the lack of balances outside the universe makes this an impractical solution, to say the least.

Another way out might be to use the most precise definition of mass provided by general relativity, the so-called *ADM mass*. However, the definition of this requires a specified behaviour at infinity, i.e., a background, which the universe lacks.

We are then left with the other general-relativistic method: determining the mass of the universe by measuring its average curvature. Let us take the defining expressions for average curvature κ for a region of size R, namely Vol. II, page 189

$$\kappa = \frac{1}{r^2_{\text{curvature}}} = \frac{3}{4\pi}\,\frac{4\pi R^2 - S}{R^4} = \frac{15}{4\pi}\,\frac{4\pi R^3/3 - V}{R^5} . \tag{111}$$

We have to insert the horizon radius R_0 and either its surface area S_0 or its volume V_0.

* In fact, at everyday energies the density of the universe lies midway between the two values, yielding the strange relation Challenge 66 e

$$m_0^2/R_0^2 \approx m_{\mathrm{Pl}}^2/R_{\mathrm{Pl}}^2 = c^4/G^2 . \tag{110}$$

Vol. V, page 145
But this fascinating relation is not new. The approximate equality can be deduced from equation 16.4.3 (p. 620) of Steven Weinberg, *Gravitation and Cosmology*, Wiley, 1972, namely $Gn_b m_p = 1/t_0^2$. The relation is required by several cosmological models.

However, given the error margins on the radius and the volume, especially at Planck energy, we again find no *reliable* result for the radius of curvature.

Challenge 68 e

Ref. 102

An equivalent method starts with the usual expression provided by Rosenfeld for the indeterminacy $\Delta\kappa$ in the scalar curvature for a region of size R, namely

$$\Delta\kappa > \frac{16\pi l_{\mathrm{Pl}}^2}{R^4} \,. \qquad (112)$$

However, this expression also shows that the error in the radius of curvature behaves like the error in the distance to the horizon.

We find that *at Planck energy, the average radius of curvature of nature lies between infinity and the Planck length*. This implies that the mass density of the universe lies between the minimum value and the Planck value. There is thus no method to determine the mass of the universe at Planck energy. (Can you find one?)

Challenge 69 s

In summary, mass measurements of the universe vary with the energy scale. Both at the lowest and at the highest energies, a precise mass value cannot be determined. The concept of mass cannot be applied to the universe as a whole: *the universe has no mass.*

Do symmetries exist in nature?

We have already seen that at the cosmological horizon, space-time translation symmetry breaks down. Let us have a quick look at the other symmetries.

What happens to permutation symmetry? Permutation is an operation on objects in space-time. It thus necessarily requires a distinction between matter, space and time. If we cannot distinguish positions, we cannot talk about exchange of particles. Therefore, at the horizon, general relativity and quantum theory together make it impossible to define permutation symmetry.

Let us explore CPT symmetry. As a result of measurement errors or of limiting maximum or minimum values, it is impossible to distinguish between the original and the transformed situations. Therefore we cannot claim that CPT is a symmetry of nature at horizon scales. In other words, matter and antimatter cannot be distinguished at the horizon.

Also gauge symmetry is not valid at horizon scale, as you may wish to check in detail yourself. For its definition, the concept of gauge field requires a distinction between time, space and mass; at the horizon this is impossible. We therefore also deduce that at the horizon, concepts such as algebras of observables cannot be used to describe nature. Renormalization breaks down too.

Challenge 70 e

All symmetries of nature break down at the cosmological horizon. None of the vocabulary we use to talk about observations – including terms such as such as 'magnetic field', 'electric field', 'potential', 'spin', 'charge', or 'speed' – can be used at the horizon.

Does the universe have a boundary?

It is common to take 'boundary' and 'horizon' as synonyms in the case of the universe, because they are the same for all practical purposes. Knowledge of mathematics does not help us here: the properties of mathematical boundaries – for example, that they themselves have no boundary – are not applicable to the universe, since space-time is not continuous. We need other, physical arguments.

The boundary of the universe is supposed to represent the boundary between *something* and *nothing*. There are three possible interpretations of 'nothing':

— 'Nothing' could mean 'no matter'. But we have just seen that this distinction cannot be made at Planck scales. So either the boundary will not exist at all or it will encompass the horizon *as well as* the whole universe.
— 'Nothing' could mean 'no space-time'. We then have to look for those domains where space and time cease to exist. These occur at Planck scales and at the horizon. Again, either the boundary will not exist or it will encompass the whole universe.
— 'Nothing' could mean 'neither space-time nor matter'. The only possibility is a boundary that encloses domains *beyond* the Planck scales and *beyond* the horizon; but again, such a boundary would also encompass all of nature.

This is puzzling. When combining quantum theory and relativity, we do not seem to be able to find a conceptual definition of the horizon that distinguishes it from what Challenge 71 s it includes. A distinction *is* possible in general relativity alone, and in quantum theory alone; but as soon as we combine the two, the boundary becomes indistinguishable from its content. *The interior of the universe cannot be distinguished from its horizon.* There is *no* boundary of the universe.

The difficulty in distinguishing the horizon from its contents suggests that nature may be *symmetric* under transformations that exchange interiors and boundaries. This idea is called *holography*, because it vaguely recalls the working of credit-card holograms. It Ref. 103 is a busy research field in high-energy physics.

We note that if the interior and the boundary of the universe cannot be distinguished, the constituents of nature can neither be points nor tiny objects of any kind. The constituents of nature must be *extended*. But before we explore this topic, we continue with our search for differences between the universe and nothing. The search leads us to our next question.

Is the universe a set?

> Domina omnium et regina ratio.*
>
> Cicero

We are used to thinking of the universe the sum of all matter and all space-time. In doing so, we imply that the universe is a set of mutually distinct components. This idea has been assumed in three situations: in claiming that matter consists of particles; that space-time consists of events (or points); and that different states consist of different initial conditions. However, our discussion shows that the universe is *not* a set of such distinguishable elements. We have encountered several proofs: at the horizon, at the big bang and at Planck scales, it becomes impossible to distinguish between events, between particles, between observables, and between space-time and matter. In those domains, distinctions of any kind become impossible. We have found that distinguishing between

* 'The mistress and queen of all things is reason.' *Tusculanae Disputationes*, 2.21.47. Marcus Tullius Cicero (106–43 BCE), was an influential lawyer, orator, writer and politician at the end of the Roman republic.

two entities – for example, between a toothpick and a mountain – is only *approximately* possible. It is approximately possible because we live at energies well below the Planck energy. The approximation is so good that we do not notice the error when we distinguish cars from people and from toothpicks. Nevertheless, our discussion of the situation at Planck energy shows that a perfect distinction is impossible in principle. *It is impossible to split the universe into separate parts.*

Another way to reach this result is the following. Distinguishing between two entities requires different measurement results: for example, different positions, masses or sizes. Whatever quantity we choose, at Planck energy the distinction becomes impossible. Only at everyday energies is it approximately possible.

In short, since the universe contains no distinguishable parts, there are *no (mathematical) elements* in nature. Simply put:

▷ The universe is not a set.

Vol. III, page 328 We envisaged this possibility earlier on; now it is confirmed. The concepts of 'element' and 'set' are already too specialized to describe the universe. *The universe must be described by a mathematical concept that does not contain any set.* The new concept must be more general than that of a set.

This is a powerful result: a precise description of the universe cannot use any concept that presupposes the existence of sets. But all the concepts we have used so far to describe nature, such as space-time, metric, phase space, Hilbert space and its generalizations, are based on elements and sets. They must all be abandoned at Planck energies, and in any precise description.

Elements and sets must be abandoned. Note that this radical conclusion is deduced from only two statements: the necessity of using quantum theory whenever the dimensions are of the order of the Compton wavelength, and of using general relativity whenever the dimensions are of the order of the Schwarzschild radius. Together, they mean that no precise description of nature can contain elements and sets. The difficulties in complying with this result explain why the unification of the two theories has not so far been successful. Not only does unification require that we stop using space, time and mass for the description of nature; it also requires that all distinctions, of any kind, should be only approximate. But all physicists have been educated on the basis of exactly the opposite creed!

Many past speculations about the unified description of nature depend on sets. In particular, all studies of quantum fluctuations, mathematical categories, posets, involved mathematical spaces, computer programs, Turing machines, Gödel's incompleteness theorem, creation of any sort, space-time lattices, quantum lattices and Bohm's unbroken wholeness *presuppose sets.* In addition, all speculations by cosmologists about the origin of the universe presuppose sets. But since these speculations presuppose sets, they are wrong. You may also wish to check the religious explanations you know against this criterion. In fact, no approach used by theoretical physicists up to the year 2000 satisfied the requirement that elements and sets must be abandoned.

Ref. 104 · Challenge 72 e

The task of abandoning sets is not easy. This is shown with a simple test: do you know of a single concept not based on elements or sets?

Challenge 73 s

In summary, the universe is not a set. In particular,

▷ The universe is not a physical system.

Specifically, the universe has no state, no intrinsic properties, no wave function, no initial conditions, no energy, no mass, no entropy and no cosmological constant. The universe is thus neither thermodynamically closed nor open; and it contains no information. All thermodynamic quantities, such as entropy, temperature and free energy, are defined using *ensembles*. Ensembles are limits of systems which are thermodynamically either open or closed. As the universe is neither open nor closed, no thermodynamic quantity can be defined for it.* All physical properties are defined only for parts of nature. Only parts of nature are approximated or idealized as sets, and thus only parts of nature are physical systems.

Curiosities and fun challenges about the universe

> Insofern sich die Sätze der Mathematik auf die Wirklichkeit beziehen, sind sie nicht sicher, und sofern sie sicher sind, beziehen sie sich nicht auf die Wirklichkeit.**
>
> Albert Einstein

> Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk.***
>
> Leopold Kronecker

In mathematics, $2 + 2 = 4$. This statement is an idealization of statements such as 'two apples plus two apples makes four apples.' However, we now know that at Planck energy, the statement about apples is not a correct statement about nature. At Planck energy, objects cannot be counted or even defined, because separation of objects is not possible at that scale. We can count objects only because we live at energies much lower than the Planck energy.

The statement by Kronecker must thus be amended. Since all integers are low-energy approximations, and since we always use low-energy approximations when talking or thinking, we provokingly conclude: man also makes the integers.

* *

If vacuum cannot be distinguished from matter or radiation, and if the universe cannot be distinguished from nothing, then it is incorrect to claim that "the universe appeared from nothing." The naive idea of creation is a logical impossibility. "Creation" results from a lack of imagination.

* *

* Some people knew this long before physicists did. For example, the belief that the universe is or contains information was ridiculed most thoroughly in the popular science-fiction parody by Douglas Adams, *The Hitchhiker's Guide to the Galaxy*, 1979, and its sequels.

** 'In so far as mathematical statements describe reality, they are not certain, and as far as they are certain, they are not a description of reality.'

*** 'Gracious god made the integers, all else is the work of man.' Leopold Kronecker (b. 1823 Liegnitz, d. 1891 Berlin) was a well-known mathematician. Among others, the Kronecker delta and the Kronecker product are named for him.

In 2002, Seth Lloyd estimated how much information the universe can contain, and how Ref. 105 many calculations it has performed since the big bang. This estimate is based on two ideas: that the number of particles in the universe is a well-defined quantity, and that the universe is a computer, i.e., a physical system. We now know that neither assumption is correct. *The universe contains no information.* Conclusions such as this one show the power of the criteria that we have deduced for any precise or complete description of motion.

* *

Astronomers regularly take pictures of the cosmic background radiation and its variations. Is it possible that these photographs will show that the spots in one direction of Challenge 74 s the sky are exactly the same as those in the diametrically opposite direction?

* *

In 1714, the famous scientist and thinker Leibniz (b. 1646 Leipzig, d. 1716 Hannover) published his *Monadologie*. In it he explores what he calls a 'simple substance', which he defined to be a substance that has no parts. He called it a *monad* and describes some Ref. 106 of its properties. However, mainly thanks to his incorrect deductions, the term has not been generally adopted. What is the physical concept most closely related to that of a Challenge 75 s monad?

* *

We usually speak of *the* universe, implying that there is only one of them. Yet there is a simple case to be made that 'universe' is an observer-dependent concept, since the idea of 'all' is observer-dependent. Does this mean that there are many universes, or a Challenge 76 s 'multiverse'?

* *

Challenge 77 e Is the 'radius' of the universe observer-invariant?

* *

Challenge 78 e Is the cosmological constant Λ observer-invariant?

* *

If all particles were removed (assuming one knew where to put them), there wouldn't be Challenge 79 s much of a universe left. True?

* *

Can you show that the distinction between matter and antimatter is not possible at the Challenge 80 e cosmic horizon? And the distinction between real and virtual particles?

* *

At Planck energy, interactions cannot be defined. Therefore, 'existence' cannot be Challenge 81 s defined. In short, at Planck energy we cannot say whether particles exist. True?

Hilbert's sixth problem settled

Vol. III, page 282
Ref. 107

In the year 1900, David Hilbert* gave a famous lecture in which he listed 23 of the great challenges facing mathematics in the twentieth century. Most of these provided challenges to many mathematicians for decades afterwards. A few are still unsolved, among them the sixth, which challenged mathematicians and physicists to find an *axiomatic* treatment of physics. The problem has remained in the minds of many physicists since that time. Scholars have developed axiomatic treatments if classical mechanics, electrodynamics and special relativity. Then they did this for quantum theory, quantum field theory and general relativity.

Whenever we combine quantum theory and general relativity, we must abandon the concept of point particle, of space point and of event. Mathematically speaking, when we combine quantum theory and general relativity, we find that nature does not contain sets, and that the universe is not a set. However, *all* mathematical systems – be they algebraic systems, order systems, topological systems or a mixture of these – are based on elements and sets. Mathematics does not have axiomatic systems without elements or sets. The reason for this is simple: every (mathematical) *concept* contains at least one element and one set. However, nature is different. And since nature does not contain sets, an axiomatic description of nature is *impossible*.

All concepts used in physics before the year 2000 depend on elements and sets. For humans, it is difficult even to *think* without first defining a set of possibilities. Yet nature does not contain sets.

▷ There is no axiomatic description of nature.

And since an axiomatic formulation of physics is impossible, we conclude that the complete, unified theory cannot be based on axioms. This is surprising at first, because separate axiomatic treatments of quantum theory and general relativity *are* possible. However, *axiomatic systems in physics are always approximate.* The need to abandon axioms is one of the reasons why reaching a unified description of nature is a challenge.

The impossibility of an axiomatic system for physics is also confirmed in another way. Physics starts with a *circular* definition: space-time and vacuum are defined with the help

Vol. I, page 437

of objects and objects are defined with the help of space-time and vacuum. In fact, physics has *never* been axiomatic! Physicists have always had to live with circular definitions.

The situation is similar to a child's description of the sky as 'made of air and clouds'. Looking closely, we discover that clouds are made up of water droplets. We find that there is air inside clouds, and that there is also water vapour away from clouds. When clouds and air are viewed through a microscope, there is no clear boundary between the two. We cannot define either of the terms 'cloud' and 'air' without the other.

Like clouds and air, also objects and vacuum are indistinguishable. Virtual particles are found in vacuum, and vacuum is found inside objects. At Planck scales there is no clear boundary between the two; we cannot define either of the terms 'particle' and 'vacuum' without the other. But despite the lack of a clean definition, and despite the

* David Hilbert (b. 1862 Königsberg, d. 1943 Göttingen) was the greatest mathematician of his time. His textbooks are still in print.

logical problems that can ensue, in both cases the description works well at large, everyday scales.

In summary, an axiomatic description of nature is impossible. In particular, the complete, unified theory must contain circular definitions. We will find out how to realize the requirement later on. Page 221

The perfect physics book

A *perfect* physics book describes *all* of nature with *full precision*. In particular, a perfect physics book describes itself, its own production, its own author, its own readers and its own contents. Can such a book exist?

Since the universe is not a set, a perfect physics book *can* exist, as it does not contradict any property of the universe. Since the universe is not a set and since it contains no information, the paradox of the perfect physics book disappears. Vol. I, page 436 Indeed, any existing physics book attempts to be perfect. But now a further question arises.

Does the universe make sense?

> Drum hab ich mich der Magie ergeben,
> [...]
> Daß ich erkenne, was die Welt
> Im Innersten zusammenhält.*
>
> Goethe, *Faust*.

Is the universe really the sum of matter–energy and space-time? Or of particles and vacuum? We have heard these statements so often that we may forget to check them. We do not need magic, as Faust thought: we only need to list what we have found so far, especially in this section, in the section on Planck scales, and in the chapter on brain and language. Vol. III, page 253 Table 3 shows the result.

Not only are we unable to state that the universe is made of space-time and matter; we are unable to say anything about the universe at all! It is not even possible to say that it exists, since it is impossible to interact with it. The term 'universe' does not allow us to make a single sensible statement. (Can you find one?) Challenge 82 r We are only able to list properties it does *not* have. We are unable to find any property that the universe *does* have. Thus, the universe has no properties! We cannot even say whether the universe is something or nothing. *The universe isn't anything in particular.* The term universe has *no* content.

By the way, there is another well-known, non-physical concept about which nothing can be said. Many scholars have explored it in detail. What is it? Challenge 83 s

In short, the term 'universe' is not at all useful for the description of motion. We can obtain a confirmation of this strange conclusion from an earlier chapter. Vol. III, page 284 There we found that any concept needs defined content, defined limits and a defined domain of application. In this section, we have found that the term 'universe' has none of these; there is thus no such concept. If somebody asks why the universe exists, the answer is: not only does the use of the word 'why' wrongly suggest that something may exist outside the universe, providing a reason for it and thus contradicting the definition of the term

* 'Thus I have devoted myself to magic, [...] that I understand how the innermost world is held together.' Goethe was a German poet.

TABLE 3 Statements about the universe when explored at highest precision, i.e., at Planck scales

The universe has no age.	The universe has no beginning.
The universe has no size.	The universe has no volume.
The universe has no shape.	The universe's particle number is undefined.
The universe has no mass.	The universe has no energy.
The universe has no density.	The universe contains no matter.
The universe has no cosmological constant.	The universe has no initial conditions.
The universe has no state.	The universe has no wave function.
The universe is not a physical system.	The universe contains no information.
The universe is not isolated.	The universe is not open.
The universe has no boundaries.	The universe does not interact.
The universe cannot be measured.	The universe cannot be said to exist.
The universe cannot be distinguished from nothing.	The universe cannot be distinguished from a single event.
The universe contains no moments.	The universe is not composite.
The universe is not a set.	The universe is not a concept.
The universe cannot be described.	There is no plural for 'universe'.
The universe cannot be distinguished from vacuum.	The universe was not created.

'universe' itself; but more importantly, the universe does not exist, because there is no such concept as a 'universe'.

In summary, any sentence containing the word 'universe' is meaningless. The word only *seems* to express something, but it doesn't.* This conclusion may be interesting, even strangely beautiful, but does it help us to understand motion more precisely? Yes, it does.

Abandoning sets and discreteness eliminates contradictions

Our discussion of the term 'universe' shows that the term cannot include any element or set. And the same applies to the term 'nature'. Nature cannot be made of atoms. Nature cannot be made of space-time points. Nature cannot be made of separate, distinct and discrete entities.

The difficulties in giving a sharp definition of 'universe' also show that the fashionable term 'multiverse' makes no sense. There is no way to define such a term, since there is no empirical way and also no logical way to distinguish 'one' universe from 'another': the universe has no boundary. In short, since the term 'universe' has no content, the term 'multiverse' has even less. The latter term has been created only to trick the media and various funding agencies. In fact, the same might even be said of the former term...

* Of course, the term 'universe' still makes sense if it is defined more restrictively: for example, as everything interacting with a particular human or animal observer in everyday life. But such a definition, equating 'universe' and 'environment', is not useful for our quest, as it lacks the precision required for a description of motion.

TABLE 4 Properties of nature at maximal, everyday and minimal scales

Physical property of nature	At horizon scale	At everyday scale	At Planck scales
requires quantum theory *and* relativity	true	false	true
intervals can be measured precisely	false	true	false
length and time intervals appear	limited	unlimited	limited
space-time is not continuous	true	false	true
points and events cannot be distinguished	true	false	true
space-time is not a manifold	true	false	true
space is 3-dimensional	false	true	false
space and time are indistinguishable	true	false	true
initial conditions make sense	false	true	false
space-time fluctuates	true	false	true
Lorentz and Poincaré symmetry	do not apply	apply	do not apply
CPT symmetry	does not apply	applies	does not apply
renormalization	does not apply	applies	does not apply
permutation symmetry	does not apply	applies	does not apply
interactions and gauge symmetries	do not exist	exist	do not exist
number of particles	undefined	defined	undefined
algebras of observables	undefined	defined	undefined
matter indistinguishable from vacuum	true	false	true
boundaries exist	false	true	false
nature is a set	false	true	false

So far, by taking into account the limits on length, time, mass and all the other quantities we have encountered, we have reached a number of almost painful conclusions about nature. However, we have also received something in exchange: all the contradictions between general relativity and quantum theory that we mentioned at the beginning of this chapter are now resolved. *We changed the contradictions to circular definitions.* Although we have had to leave many cherished habits behind us, in exchange we have the promise of a description of nature without contradictions. But we get even more.

Challenge 84 e

Extremal scales and open questions in physics

Page 19

At the beginning of this volume, we listed all the fundamental properties of nature that are unexplained either by general relativity or by quantum theory. We called it the millennium list. The results of this chapter provide us with surprising statements on many of the items. In fact, many of the statements are not new at all, but are surprisingly familiar. Let us compare systematically the statements from this chapter, on the universe, with those of the previous chapter, on Planck scales. The comparison is given in Table 4.

First, Table 4 shows that *each* unexplained property listed there is unexplained at *both* limits of nature, the small and the large limit. Worse, many of these unexplained general properties do not even *make sense* at the two limits of nature!

Secondly, and more importantly, nature behaves in the *same way* at the cosmological horizon scale and at the Planck scale. In fact, we have not found any difference between
Challenge 85 r
the two cases. (Can you discover one?) We are thus led to the hypothesis that nature does not distinguish between the large and the small. Nature seems to be characterized by *extremal identity*.

Is extremal identity a principle of nature?

The idea of extremal identity incorporates some rather general points:

— All open questions about nature appear at both size extremes.
— Any description of nature requires both general relativity and quantum theory.
— Nature, or the universe, is not a set.
— Initial conditions and evolution equations make no sense at nature's limits.
— There is a relation between local and global issues in nature.
— The concept of 'universe' has no content.

Extremal identity thus looks like a useful hypothesis in the search for a unified description of nature. To be a bit more provocative, it seems that extremal identity may be the *only* hypothesis incorporating the idea that the universe is not a set. Therefore, extremal identity seems to be essential in the quest for unification.

Extremal identity is beautiful in its simplicity, in its unexpectedness and in the rich-
Challenge 86 e
ness of its consequences. You might enjoy exploring it by yourself. In fact, the explora-
Ref. 108
tion of extremal identity is currently the subject of much activity in theoretical physics, although often under different names.

The simplest approach to extremal identity – in fact, one that is too simple to be correct – is *inversion*. Indeed, extremal identity seems to imply a connection such as

$$r \leftrightarrow \frac{l_{\mathrm{Pl}}^2}{r} \quad \text{or} \quad x_\mu \leftrightarrow \frac{l_{\mathrm{Pl}}^2\, x_\mu}{x_\mu\, x^\mu} \tag{113}$$

relating distances r or coordinates x_μ with their inverse values using the Planck length l_{Pl}. Can this mapping be a symmetry of nature? At every point of space? For example, if the horizon distance is inserted, the relation (113) implies that lengths smaller than
Challenge 87 s
$l_{\mathrm{Pl}}/10^{61} \approx 10^{-96}$ m never appear in physics. Is this the case? What would inversion imply for the big bang?

More involved approaches to extremal identity come under the name of *space-time*
Ref. 103
duality and *holography*. They are subject of intense research. Numerous fascinating questions are contained in extremal identity; there is a lot of fun ahead of us.

Above all, we need to find the correct version of the inversion relation (113). Inversion is neither sufficient nor correct. It is not sufficient because it does not explain *any* of the millennium issues left open by general relativity and quantum theory. It only *relates* some of them, but it does not *solve* any of them. (You may wish to check this for yourself.)
Challenge 88 e
In other words, we need to find the precise description of quantum geometry and of elementary particles.

However, inversion is also simply wrong. Inversion is not the correct description of extremal identity because it does not realize a central result discovered above: it does not connect *states* and *intrinsic properties*, but keeps them distinct. In particular, inversion does not take *interactions* into account. And most open issues at this point of our mountain ascent concern the properties and the appearance of interactions. Page 86

SUMMARY ON THE UNIVERSE

The exploration of the universe allows us to formulate several additional requirements for the complete theory that we are looking for:

— Whenever we combine general relativity and quantum theory, the universe teaches us that *it is not a set of parts*. For this reason, any sentence or expression containing the term 'universe' is meaningless whenever *full precision* is required.
— We learned that a description of nature without sets *solves the contradictions* between general relativity and quantum theory.
— We found, again, that despite the contradictions between quantum theory and general relativity, the Planck limits c, $\hbar$ and $c^4/4G$ *remain valid*.
— We then found an intriguing relation between Planck scales and cosmological scales: they seem to pose the same challenges to their description. There is a close relationship between large and small scales in nature.

We can now answer the question in the chapter title: there seems to be little difference – if any at all – between the universe and nothing. We can express this result in the following catchy statement:

▷ The universe cannot be observed.

In our journey, the confusion and tension are increasing. But in fact we are getting close to our goal, and it is worth continuing.

A PHYSICAL APHORISM

Here is a humorous 'proof' that we really are near the top of Motion Mountain. Salecker and Wigner, and then Zimmerman, formulated the fundamental limit for the measurement precision τ attainable by a clock of mass M. It is given by $\tau = \sqrt{\hbar T/c^2 M}$, where T is the time to be measured. We can then ask what maximum time T can be measured with a precision of a Planck time $t_{\rm Pl}$, given a clock of the mass of the whole universe. We get a maximum time of Ref. 55, Ref. 56

$$T = \frac{t_{\rm Pl}^2 c^2}{\hbar} M \,. \tag{114}$$

Inserting numbers, we find rather precisely that the time T is the present age of the universe. Challenge 89 s

With the right dose of humour we can see this result as a sign that time is now ripe, after so much waiting, for us to understand the universe down to the Planck scales. We are thus getting nearer to the top of Motion Mountain. Be prepared for a lot of fun.

Chapter 6

THE SHAPE OF POINTS – EXTENSION IN NATURE

> “Nil tam difficile est, quin quaerendo investigari possiet.**”
>
> Terence

The usual expressions for the reduced Compton wavelength $\lambda\!\!\!^{-} = \hbar/mc$ and for the Schwarzschild radius $r_s = 2Gm/c^2$, taken together, imply the conclusion that at Planck energies, what we call ‘space points’ and ‘point particles’ must actually be described by *extended* constituents that are infinite and fluctuating in size. We will show this result with the following arguments:

1. Any experiment trying to measure the size or the shape of an elementary particle with high precision inevitably leads to the result that at least one dimension of the particle is of macroscopic size.
2. There is no evidence that empty space consists of points, as they cannot be measured or detected. In addition, in order to build up a measurable entity, such as the vacuum, that is extended in three dimensions, its constituents must also be extended.
3. The existence of minimum measurable distances and time intervals implies the existence of space-time duality: a symmetry between very large and very small distances. Space-time duality in turn implies that the fundamental constituents that make up vacuum and matter are extended.
4. The constituents of the universe, and thus of vacuum, matter and radiation, cannot form a (mathematical) set. And any precise description of nature without sets must use extended constituents.
5. The Bekenstein–Hawking expression for the entropy of black holes – in particular its surface dependence – confirms that both vacuum and particles are composed of extended constituents.
6. The attempt to extend statistical properties to Planck scales shows that both particles and space points behave as extended constituents at high energies, namely as braids or tangles.
7. The belt trick provides a model for fermions that matches observations and again suggests extended constituents in matter.

We conclude the chapter with some experimental and theoretical checks of extension

** ‘Nothing is so difficult that it could not be investigated.’ Terence is Publius Terentius Afer (b. *c.* 190 Carthago, d. 159 BCE Greece), important Roman poet. He writes this in his play *Heauton Timorumenos*, verse 675.

and an overview of present research efforts.

> Also, die Aufgabe ist nicht zu sehen, was noch nie jemand gesehen hat, sondern über dasjenige was jeder schon gesehen hat zu denken was noch nie jemand gedacht hat.*
>
> Erwin Schrödinger

The size and shape of elementary particles

Size is the length of vacuum taken by an object. This definition comes naturally in everyday life, quantum theory and relativity. To measure the size of an object as small as an elementary particle, we need high energy. The higher the energy, the higher the precision with which we can measure the size.

Page 89 However, near the Planck energy, vacuum and matter cannot be distinguished: it is impossible to define the boundary between the two, and thus it is impossible to define the size of an object. As a consequence, every object, and in particular every elementary particle, becomes as extended as the vacuum! There is no measurement precision at all at Planck scales. Can we save the situation? Let us take a step back. Do measurements at least allow us to say whether particles can be contained inside small spheres?

Do boxes exist?

The first and simplest way to determine the size of a compact particle such as a sphere, or to find at least an upper limit, is to measure the size of a *box* it fits in. To be sure that the particle is inside, we must first be sure that the box is *tight*: that is, whether anything (such as matter or radiation) can leave the box.

But there is no way to ensure that a box has no holes! We know from quantum physics that any wall is a *finite* potential hill, and that tunnelling is always possible. In short, there is no way to make a completely tight box.

Let us cross-check this result. In everyday life, we call particles 'small' when they can be enclosed. Enclosure is possible in daily life because walls are impenetrable. But walls are only impenetrable for matter particles up to about 10 MeV and for photons up to about 10 keV. In fact, boxes do not even exist at medium energies. So we certainly cannot extend the idea of 'box' to Planck energy.

Since we cannot conclude that particles are of compact size by using boxes, we need to try other methods.

Can the Greeks help? – The limitations of knives

The Greeks deduced the existence of atoms by noting that matter cannot be divided indefinitely. There must be *uncuttable* particles, which they called *atoms*. Twenty-five cen-

* 'Our task is not to see what nobody has ever seen, but to think what nobody has ever thought about that which everybody has seen already.' Erwin Schrödinger (b. 1887 Vienna, d. 1961 Vienna) discovered the equation describing non-relativistic quantum motion; it brought him international fame and the Nobel Prize in Physics.

turies later, experiments in the field of quantum physics confirmed the conclusion, but modified it: nowadays, the elementary particles are the 'atoms' of matter and radiation.

Despite the huge success of the concept of elementary particle, at Planck energy, we have a different situation. The use of a knife, like any other cutting process, is the insertion of a wall. Walls and knives are potential hills. All potential hills are of finite height, and allow tunnelling. Therefore a wall is never perfect, and thus neither is a knife. In short, any attempt to divide matter fails to work when we approach Planck scales. At Planck energy, any subdivision is impossible.

The limitations of knives and walls imply that at Planck energy, an attempted cut does not necessarily lead to two separate parts. At Planck energy, we can never state that the two parts have been really, completely separated: the possibility of a thin connection between the two parts to the right and left of the blade can never be excluded. In short, at Planck scales we cannot prove compactness by cutting objects.

Are cross sections finite?

To sum up: despite all attempts, we cannot show that elementary particles are point-like. Are they, at least, of finite size?

To determine the size of a particle, we can try to determine its departure from point-likeness. Detecting this departure requires scattering. For example, we can suspend the particle in a trap and then shoot a probe at it. What happens in a scattering experiment at

Ref. 109 highest energies? This question has been studied by Leonard Susskind and his colleagues. When shooting at the particle with a high-energy probe, the scattering process is characterized by an interaction time. Extremely short interaction times imply sensitivity to size and shape fluctuations, due to the quantum of action. An extremely short interaction time also provides a cut-off for high-energy shape and size fluctuations, and thus determines the measured size. As a result, the size measured for any microscopic, but extended, object *increases* when the probe energy is increased towards the Planck value.

In summary, even though at experimentally achievable energies the size of an elementary particle is always smaller than the measurement limit, when we approach the Planck energy, the particle size increases beyond all bounds. So at high energies we cannot give an upper limit to the size of a particle – except the universe itself. In other words, since particles are not point-like at everyday energies, at Planck energy they are enormous:

▷ Quantum particles are extended.

That is quite a statement. Are particles really not of finite, bounded size? Right at the

Vol. I, page 25 start of our mountain ascent, we distinguished objects from their environment. Objects are by definition localized, bounded and compact. All objects have a *boundary*, i.e., a surface which does not itself have a boundary. Objects are also bounded in abstract ways: also the set of symmetries of an object, such as its geometric symmetry group or its gauge group, is bounded. In contrast, the environment is not localized, but extended and unbounded. But all these basic assumptions fail us at Planck scales. At Planck energy, it is impossible to determine whether something is bounded or compact. Compactness and locality are only approximate properties; they are not applicable at high energies.

TABLE 5 Effects of various camera shutter times on photographs

Duration	Blur	Observation possibilities and effects
1 h	high	Ability to see faint quasars at night if motion is compensated
1 s	high	Everyday motion completely blurred
20 ms	lower	Interruption by eyelids; small changes impossible to see
10 ms	lower	Effective eye/brain shutter time; tennis ball impossible to see while hitting it
0.25 ms	lower	Shortest commercial photographic camera shutter time; ability to photograph fast cars
1 µs	very low	Ability to photograph flying bullets; strong flashlight required
c. 10 ps	lowest	Study of molecular processes; ability to photograph flying light pulses; laser light required to get sufficient illumination
10 fs	higher	Light photography impossible because of wave effects
100 zs	high	X-ray photography impossible; only γ-ray imaging left over
shorter times	very high	Photographs get darker as illumination decreases; gravitational effects significant
10^{-43} s	highest	Imaging impossible

In particular, the idea of a point particle is an approximate concept, valid only at low energies.

We conclude that particles at Planck scales are as extended as the vacuum. Let us perform another check.

Can we take a photograph of a point?

> “Καιρὸν γνῶθι.*”
>
> Pittacus

Humans – or any other types of observers – can only observe the world with finite resolution in time and in space. In this respect, humans resemble a film camera. Every camera has a resolution limit: it can only distinguish two events if they are a certain minimum distance apart and separated by a certain minimum time. What is the best resolution possible? The value was (almost) discovered in 1899: the Planck time and the Planck length. No human, no film camera and no apparatus can measure space or time intervals smaller than the Planck values. But what would happen if we took photographs with shutter times that *approach* the Planck time?

Ref. 110

Ref. 47, Ref. 24

Imagine that you have the world's best shutter and that you are taking photographs at shorter and shorter times. Table 5 gives a rough overview of the possibilities. When shutter times are shortened, photographs get darker and sharper. When the shutter time reaches the oscillation time of light, strange things happen: light has no chance to pass undisturbed; signal and noise become indistinguishable; and the moving shutter will

* 'Recognize the right moment.' Also rendered as: 'Recognize thy opportunity.' Pittacus (Πιττακος) of Mytilene (*c.* 650–570 BCE), was a Lesbian tyrant and lawmaker; he was also one of the 'Seven Sages' of ancient Greece.

produce colour shifts. In contrast to our everyday experience, the photograph would get more *blurred* and *incorrect* at extremely short shutter times. Photography is impossible not only at long but also at short shutter times.

The difficulty of taking photographs is independent of the wavelength used. The limits move, but do not disappear. With a shutter time of τ, photons of energy lower than $\hbar/\tau$ cannot pass the shutter undisturbed.

In short, the blur decreases when shutter times usual in everyday life are shortened, but *increases* when shutter times are shortened further towards Planck times. As a result, there is no way to detect or confirm the existence of point objects by taking pictures. Points in space, as well as instants of time, are *imagined* concepts: they do not belong in a precise description of nature.

At Planck shutter times, only signals with Planck energy can pass through the shutter. Since at these energies matter cannot be distinguished from radiation or from empty space, all objects, light and vacuum look the same. It is impossible to say what nature looks like at very short times.

But the situation is worse than this: a Planck shutter cannot exist at all, as it would need to be as small as a Planck length. A camera using it could not be built, as lenses do not work at this energy. Not even a camera obscura – without any lens – would work, as diffraction effects would make image production impossible. In other words, the idea that at short shutter times a photograph of nature shows a frozen image of everyday life, like a stopped film, is completely wrong. In fact, a shutter does not exist even at medium energy: shutters, like walls, stop existing at around 10 MeV. At a single instant of time, nature is not frozen at all. Zeno criticized this idea in his discussions of motion, though not as clearly as we can do now. At short times, nature is blurred. In particular, point particles do not exist.

In summary, whatever the intrinsic shape of what we call a 'point' might be, we know that, being always blurred, it is first of all a cloud. Whatever method is used to photograph an elementary particle, the picture is always extended. Therefore we need to study its shape in more detail.

What is the shape of an electron?

Since particles are not point-like, they have a shape. How can we determine it? We determine the shape of an everyday object by *touching* it from all sides. This works with plants, people or machines. It even works with molecules, such as water molecules. We can put them (almost) at rest, for example in ice, and then scatter small particles off them. Scattering is just a higher-energy version of touching. However, scattering cannot determine shapes of objects smaller than the wavelength of the probes used. To determine the shape of an object as small as an electron, we need the highest energies available. But we already know what happens when approaching Planck scales: the shape of a particle becomes the shape of all the space surrounding it. In short, the shape of an electron cannot be determined in this way.

Another way to determine the shape is to build a tight box around the system under investigation and fill it with molten wax. We then let the wax cool and observe the hollow part. However, near Planck energy, boxes do not exist. We are unable to determine the shape in this way.

A third way to measure the shape of an object is to cut it into pieces and then study the pieces. As is well known, the term 'atom' just means 'uncuttable' or 'indivisible'. However, neither atoms nor indivisible particles can exist. Indeed, *cutting* is just a low-energy version of a scattering process. And the process does not work at high energies. Therefore, there is no way to prove that an object is indivisible at Planck scales. Our everyday intuition leads us completely astray at Planck energy.

We could try to distinguish transverse and longitudinal shape, with respect to the direction of motion. However, for transverse shape we get the same issues as for scattering; transverse shape diverges for high energy. And to determine longitudinal shape, we need at least two infinitely high potential walls. We already know that this is impossible.

A further, indirect way of measuring shapes is to measure the moment of inertia. A finite moment of inertia means a compact, finite shape. But when the measurement energy is increased towards Planck scales, rotation, linear motion and exchange become mixed up. We do not get meaningful results. Ref. 47

Yet another way to determine shapes is to measure the *entropy* of a collection of particles we want to study. This allows us to determine the dimensionality and the number of internal degrees of freedom. But at high energies, a collection of electrons would become a black hole. We will study this issue separately below, but again we find no new information.

Are these arguments watertight? We assumed three dimensions at all scales, and that the shape of the particle itself is fixed. Maybe these assumptions are not valid at Planck scales? Let us check the alternatives. We have already shown that because of the fundamental measurement limits, the dimensionality of space-time cannot be determined at Planck scales. Even if we could build perfect three-dimensional boxes, holes could remain in other dimensions. It does not take long to see that all the arguments against compactness work even if space-time has additional dimensions. Ref. 47 Challenge 90 e

Is the shape of an electron fixed?

Only an object composed of localized constituents, such as a house or a molecule, can have a fixed shape. The smaller the system, the more quantum fluctuations play a role. No small entity of finite size – in particular, no elementary particle – can have a fixed shape. In every thought experiment involving a finite shape, the shape itself fluctuates. But we can say more.

The distinction between particles and environment rests on the idea that particles have *intrinsic* properties. In fact, all intrinsic properties, such as spin, mass, charge, and parity, are localized. But we have seen that no intrinsic property is measurable or definable at Planck scales. Thus it is impossible to distinguish particles from the environment. In addition, at Planck energy particles have all the properties that the environment has. In particular, particles are extended. Page 86

In short, we cannot prove by experiments that at Planck energy elementary particles are finite in size in all directions. In fact, all experiments we can think of are compatible with extended particles, with 'infinite' size. More precisely, a particle always reaches the borders of the region of space-time under exploration. In simple words, we can also say that particles have *tails* or *tethers*.

Not only are particles extended, but their shape cannot be determined by the methods just explored. The only remaining possibility is that suggested by quantum theory:

▷ The shape of a particle fluctuates.

We reach the same conclusion for quanta of radiation: the box argument shows that also radiation particles are extended and fluctuating.

Incidentally, we have also settled an important question about *elementary* particles. We have already seen that any particle that is smaller than its own Compton wavelength Vol. IV, page 108 must be elementary. If it were composite, there would be a lighter component inside it; this lighter particle would have a larger Compton wavelength than the composite particle. This is impossible, since the size of a composite particle must be larger than the Compton wavelength of its components.*

However, an elementary particle *can* have constituents, provided that they are not compact. The difficulties of compact constituents were described by Andrei Sakharov Ref. 28 in the 1960s. If the constituents are extended, the previous argument does not apply, as extended constituents have no localized mass. As a result, if a flying arrow – Zeno's famous example – is made of extended constituents, it cannot be said to be at a given position at a given time. Shortening the observation time towards the Planck time makes an arrow disappear in the cloud that makes up space-time.**

Summary of the first argument for extension

Point particles do not exist at Planck scales. At Planck scales, all thought experiments with particles suggest that matter and radiation are made of *extended and fluctuating constituents of infinite size*.

Ref. 47 For extended constituents, the requirement of a non-local description is satisfied. The argument forbidding composition of elementary particles is circumvented, as extended constituents have no mass. Thus the concept of Compton wavelength cannot be defined or applied to extended constituents, and elementary particles can have constituents if these constituents are extended and massless. However, if the constituents are infinitely Challenge 91 e extended, how can compact, point-like particles be formed from them? We will look at a few options shortly.

The shape of points in vacuum

> “Thus, since there is an impossibility that [finite] quantities are built from contacts and points, it is necessary that there be indivisible material elements and [finite] quantities.”
>
> Ref. 112 Aristotle,*** *Of Generation and Corruption*.

* Examples are the neutron, positronium, or the atoms. Note that the argument does not change when the elementary particle itself is unstable, like the muon. The possibility that all components are heavier than the composite, which would avoid this argument, does not seem to lead to satisfying physical properties: for example, it leads to intrinsically unstable composites.

Ref. 111 ** Thus at Planck scales there is no quantum Zeno effect.

*** Aristotle (b. 384/3 Stageira, d. 322 BCE Chalkis), Greek philosopher and scientist.

We are used to the idea that empty space is made of spatial points. However, at Planck scales, no measurement can give zero length, zero mass, zero area or zero volume. There is no way to state that something in nature is a point without contradicting experimental results.

Furthermore, the idea of a point is an extrapolation of what is found in small empty boxes getting smaller and smaller. But we have just seen that at high energies small boxes cannot be said to be empty. In fact, boxes do not exist at all, as they can never have impenetrable walls at high energies.

Also, the idea of a point as a limiting subdivision of empty space is untenable. At small distances, space cannot be subdivided, as division requires some sort of dividing wall, which is impossible.

Even the idea of repeatedly putting a point between two others cannot be applied. At high energy, it is impossible to say whether a point is exactly on the line connecting the outer two points; and near Planck energy, there is no way to find a point between them at all. In fact, the term 'in between' makes no sense at Planck scales.

We thus find that space points do not exist, just as point particles do not exist. But there are other reasons why space cannot be made of points. In order to form space. points need to be kept *apart* somehow. Indeed, mathematicians have a strong argument for why physical space cannot be made of mathematical points: the properties of mathematical spaces described by the Banach–Tarski paradox are quite different from those of the physical vacuum. The Banach–Tarski paradox states that a sphere made of mathematical points can be cut into five pieces which can be reassembled into two spheres each of the same volume as the original sphere. Mathematically, there are sets of points for which the concept of volume makes no sense. Physically speaking, we conclude that the concept of volume does not exist for continuous space; it is only definable if an *intrinsic length* exists. And in nature, an intrinsic length exists for matter and for vacuum: the Planck length. And any concept with an intrinsic length must be described by one or several extended constituents.* In summary, in order to build up space, we need *extended* constituents.

Ref. 47

Also the number of space dimensions is problematic. Mathematically, it is impossible to define the dimension of a set of points on the basis of the set structure alone. Any compact one-dimensional set has many points as any compact three-dimensional set – indeed, as any compact set of any dimensionality greater than zero. To build up the *physical* three-dimensional vacuum, we need constituents that *organize* their neighbourhood. The fundamental constituents must possess some sort of ability to form bonds, which will construct or fill precisely three dimensions. Bonds require extended constituents. A collection of tangled constituents extending to the maximum scale of the region under consideration would work perfectly. Of course, the precise shape of the fundamental constituents is not yet known. In any case, we again find that any constituents of physical three-dimensional space must be *extended*.

* Imagining the vacuum as a collection of compact constituents, such as spheres, with Planck size in all directions would avoid the Banach–Tarski paradox, but would not allow us to deduce the number of dimensions of space and time. It would also contradict all the other results of this section. Therefore we do not explore it further.

Ref. 113
Challenge 92 s

In summary, we need extension to define dimensionality and to define volume. This result is not surprising. We deduced above that the constituents of particles are extended. Since vacuum is not distinguishable from matter, we would expect the constituents of vacuum to be extended as well. Stated simply, if elementary particles are not point-like, then points in the vacuum cannot be either.

Measuring the void

To check whether the constituents of the vacuum are extended, let us perform a few additional thought experiments. First, let us measure the size of a point in space. The clearest definition of size is in terms of the cross section. How can we determine the cross section of a point? We can determine the cross section of a piece of vacuum and then determine the number of points inside it. However, at Planck energy, we get a simple result: the cross section of a volume of empty space is independent of depth. At Planck energy, vacuum has a surface, but no depth. In other words, at Planck energy we can only state that a Planck layer covers the surface of a region. We cannot say anything about its interior. One way to picture this result is to say that what we call 'space points' are in fact long tubes.

Another way to determine the size of a point is to count the points found in a given volume of space-time. One approach is to count the possible positions of a point particle in a volume. However, at Planck energy point particles are extended and indistinguishable from vacuum. At Planck energy, the number of points is given by the surface area of the volume divided by the Planck area. Again, the surface dependence suggests that particles and the constituents of space are long tubes.

What is the maximum number of particles that fit inside a piece of vacuum?

Another approach to counting the number of points in a volume is to fill a piece of vacuum with point particles.

The maximum mass that fits into a piece of vacuum is a black hole. But in this case too, the maximum mass depends only on the *surface* of the given region of vacuum. The maximum mass increases less rapidly than the volume. In other words, the number of physical points inside a region of space is only proportional to the surface area of the region. We are forced to conclude that vacuum must be made of extended constituents crossing the whole region, independently of its shape.

Summary of the second argument for extension

Planck scales imply that *space is made of fluctuating extended constituents of huge size.* Like particles, also space and vacuum are not made of points, but of a web. Vacuum requires a statistical description.

Vol. I, page 339 More than two thousand years ago, the Greeks argued that matter must be made of
Ref. 114 particles because salt can be dissolved in water and because fish can swim through water. Now that we know more about Planck scales, we have to reconsider this argument. Like fish swimming through water, particles can move through vacuum; but since vacuum has no bounds and cannot be distinguished from matter, vacuum cannot be made of

localised particles. However, another possibility allows for motion of particles through a vacuum: *both* vacuum *and* particles might be made of a web of extended constituents. Let us study this possibility in more detail.

THE LARGE, THE SMALL AND THEIR CONNECTION

> “I could be bounded in a nutshell and count myself a king of infinite space, were it not that I have bad dreams.”
>
> William Shakespeare,* *Hamlet*.

If two observables cannot be distinguished, there is a symmetry transformation connecting them. For example, by a change of observation frame, an electric field may (partially) change into a magnetic one. A symmetry transformation means that we can change the viewpoint (i.e., the frame of observation) in such a way that the same observation is described by one quantity from one viewpoint and by the corresponding quantity from the other viewpoint.

When measuring a length at Planck scales it is impossible to say whether we are measuring the length of a piece of vacuum, the Compton wavelength of a body, or the Schwarzschild diameter of a body. For example, the maximum size for an elementary object is its Compton wavelength. The minimum size for an elementary object is its Schwarzschild radius. The actual size of an elementary object is somewhere in between. If we want to measure the size precisely, we have to go to Planck energy; but then all these quantities are the same. In other words, at Planck scales, there is a symmetry transformation between Compton wavelength and Schwarzschild radius. In short, *at Planck scales there is a symmetry between mass and inverse mass.*

As a further consequence, at Planck scales there is a symmetry between size and inverse size. Matter–vacuum indistinguishability means that there is a symmetry between length and inverse length at Planck energy. This symmetry is called *space-time duality* or *T-duality* in the research literature of superstrings.** Space-time duality is a symmetry between situations at scale $n\,l_{\mathrm{Pl}}$ and at scale $f l_{\mathrm{Pl}}/n$, or, in other words, between R and $(f l_{\mathrm{Pl}})^2/R$, where the number f is usually conjectured to have a value somewhere between 1 and 1000. Ref. 115

Duality is a genuine non-perturbative effect. It does not exist at low energy, since duality automatically also relates energies E and $E_{\mathrm{Pl}}^2/E = \hbar c^3/GE$, i.e., it relates energies below and above Planck scale. Duality is not evident in everyday life. It is a quantum symmetry, as it includes Planck's constant in its definition. It is also a general-relativistic effect, as it includes the gravitational constant and the speed of light. Let us study duality in more detail.

* William Shakespeare (1564 Stratford upon Avon–1616 Stratford upon Avon) wrote theatre plays that are treasures of world literature.

** There is also an *S-duality*, which connects large and small coupling constants, and a *U-duality*, which is the combination of S- and T-duality.

Is small large?

> [Zeno of Elea maintained:] If the existing are many, it is necessary that they are at the same time small and large, so small to have no size, and so large to be without limits.
>
> Simplicius*

Ref. 116

To explore the consequences of duality, we can compare it to rotational symmetry in everyday life. Every object in daily life is symmetrical under a full rotation of 2π. For the rotation of an observer, angles make sense only as long as they are smaller than 2π. If a rotating observer were to insist on distinguishing angles of $0, 2\pi, 4\pi$ etc., he would get a new copy of the universe at each full turn.

Similarly, in nature, scales R and $l_{\rm Pl}^2/R$ cannot be distinguished. Lengths make no sense when they are smaller than $l_{\rm Pl}$. If, however, we insist on using even smaller values and on distinguishing them from large ones, we get a new copy of the universe at those small scales. Such an insistence is part of the standard continuum description of motion, where it is assumed that space and time are described by the real numbers, which are defined over arbitrarily small intervals. Whenever the (approximate) continuum description with infinite extension is used, the $R \leftrightarrow l_{\rm Pl}^2/R$ symmetry pops up.

Duality implies that diffeomorphism invariance is only valid at medium scales, not at extremal ones. At extremal scales, quantum theory has to be taken into account in the proper manner. We do not yet know how to do this.

Space-time duality means that introducing lengths smaller than the Planck length (as when one defines space points, which have size zero) means at the same time introducing things with very large ('infinite') value. Space-time duality means that for every small enough sphere the inside equals the outside.

Duality means that if a system has a small dimension, it also has a large one, and vice versa. There are thus no small objects in nature. So space-time duality is consistent with the idea that the basic constituents are extended.

Unification and total symmetry

So far, we have shown that at Planck energy, time and length cannot be distinguished, and that vacuum and matter cannot be distinguished. Duality shows that mass and inverse mass cannot be distinguished. As a consequence, we deduce that length, time, and mass cannot be distinguished from each other at *all* energies and scales! And since every observable is a combination of length, mass and time, *space-time duality means that there is a symmetry between all observables.* We call it the *total symmetry*.**

Total symmetry implies that there are many specific types of duality, one for each pair of quantities under investigation. Indeed, the number of duality types discovered is increasing every year. It includes, among others, the famous electric–magnetic duality we

Vol. III, page 91

* Simplicius of Cilicia (*c.* 499 – 560), neoplatonist philosopher.

** A symmetry between size and Schwarzschild radius, i.e., a symmetry between length and mass, leads to general relativity. Additionally, at Planck energy there is a symmetry between size and Compton wavelength. In other words, there is a symmetry between length and inverse mass. This implies a symmetry between coordinates and wave functions, i.e., a symmetry between states and observables. It leads to quantum theory.

Ref. 117 first encountered in electrodynamics, coupling constant duality, surface–volume duality, space-time duality, and many more. All this confirms that there is an enormous amount of symmetry at Planck scales. In fact, similar symmetries have been known right from Ref. 115 the beginning of research in quantum gravity.

Most importantly, total symmetry implies that gravity can be seen as equivalent to all other forces. Space-time duality thus shows that unification is possible. Physicists have always dreamt about unification. Duality tells us that this dream can indeed be realized.

It may seem that total symmetry completely contradicts what was said in the previous section, where we argued that all symmetries are lost at Planck scales. Which result is correct? Obviously, both of them are.

At Planck scales, all low-energy symmetries are indeed lost. In fact, all symmetries that imply a *fixed* energy are lost. However, duality and its generalizations combine both small and large dimensions, or large and small energies. Most of the standard symmetries of physics, such as gauge, permutation and space-time symmetries, are valid at each fixed energy separately. But nature is not made this way. The precise description of nature requires us to take into consideration large and small energies at the same time. In everyday life, we do not do that. The physics of everyday life is an approximation to nature valid at low and fixed energies. For most of the twentieth century, physicists tried to reach higher and higher energies. We believed that precision increases with increasing energy. But when we combine quantum theory and gravity we are forced to change this approach. To achieve high precision, we must take high and low energy into account at the same time.*

The great differences between the phenomena that occur at low and high energies are the main reason why unification is so difficult. We are used to dividing nature along a scale of energies: high-energy physics, atomic physics, chemistry, biology, and so on. But we are not allowed to think in this way any more. We have to take all energies into account at the same time. That is not easy, but we do not have to despair. Important conceptual progress was made in the last decade of the twentieth century. In particular, we now know that we need only *one constituent* for all things that can be measured.

Since there is only one constituent, total symmetry is automatically satisfied. And since there is only one constituent, there are many ways to study it. We can start from any (low-energy) concept in physics and explore how it looks and behaves when we approach Planck scales. In the present section, we are looking at the concept of 'point'. Obviously, the conclusions must be the same whatever concept we start with, be it electric field, spin, Challenge 93 d or any other. Such studies thus provide a check for the results in this section.

Summary of the third argument for extension

Unification implies thinking in terms of duality and the concepts that follow from it. The large and the small are connected. Duality points to one single type of extended constituents that defines *all* physical observables.

We still need to understand exactly what happens to duality when we restrict ourselves Challenge 94 e to low energies, as we do in everyday life. We explore this now.

* Renormalization energy does connect different energies, but not in the correct way; in particular, it does not include duality.

DOES NATURE HAVE PARTS?

> Pluralitas non est ponenda sine necessitate.*
> William of Occam

Ref. 118
Another argument, independent of those given so far, points towards a model of nature based on extended constituents. We know that any concept for which we can distinguish parts is described by a set. We usually describe nature as a set of objects, positions, instants and so on. The most famous set-theoretic description of nature is the oldest known, given by Democritus:

The world is made of indivisible particles and void.

This description was extremely successful in the past: there are no discrepancies with observations. However, after 2500 years, the conceptual difficulties of this approach are obvious.

We know that Democritus was wrong, first of all, because vacuum and matter cannot be distinguished at Planck scales. Thus the word 'and' in his sentence is already a mistake. Secondly, because of the existence of minimal scales, the void cannot be made of 'points', as we usually assume. Thirdly, the description fails because particles are not compact objects. Finally, total symmetry implies that we cannot distinguish parts in nature. Nothing can be distinguished from anything else with complete precision, and thus the particles or points in space that make up the naive model of the world cannot exist.

In summary, quantum theory and general relativity together show that in nature, *all partitions and all differences are only approximate*. Nothing can really be distinguished from anything else with complete precision. In other words, there is no way to define a 'part' of nature, whether for matter, space, time, or radiation.

▷ Nature cannot be a set.

Vol. III, page 328
The conclusion does not come as a surprise. We have already encountered another reason to doubt that nature is a set. Whatever definition we use for the term 'particle', Democritus cannot be correct for a purely logical reason. The description he provided is *not complete*. Every description of nature that defines nature as a set of parts fails to explain the *number* of these parts. In particular, the number of particles and the number of dimensions of space-time must be specified if we describe nature as made from particles and vacuum. For example, we saw that it is rather dangerous to make fun of the famous statement by Arthur Eddington

* 'Multitude should not be introduced without necessity.' This famous principle is commonly called *Occam's razor*. William of Ockham (b. 1285/1295 Ockham, d. 1349/50 Munich), or Occam in the common Latin spelling, was one of the great thinkers of his time. In his famous statement he expresses that only those concepts which are strictly necessary should be introduced to explain observations. It can be seen as the requirement to abandon *beliefs* when talking about nature. But at this stage of our mountain ascent it has an even more direct interpretation: the existence of *any* multitude in nature is questionable.

Ref. 119

> I believe there are 15,747,724,136,275,002,577,605,653,961,181,555,468,044, 717,914,527,116,709,366,231,425,076,185,631,031,296 protons in the universe and the same number of electrons.

In fact, practically all physicists share this belief, although they usually either pretend to favour some other number, or worse, keep the number unspecified.

In modern physics, many specialized sets are used to describe nature. We have used vector spaces, linear spaces, topological spaces and Hilbert spaces. But so far, we consistently refrained, like all physicists, from asking about the origin of their sizes (mathematically speaking, of their dimensionality or cardinality). In fact, it is just as unsatisfying to say that the universe contains some specific number of atoms as it is to say that space-time is made of point-like events arranged in $3+1$ dimensions. Both are statements about set sizes, in the widest sense. In a complete, unified description of nature the number of smallest particles and the number of space-time points must not be fixed beforehand, but must *result* from the description.

Any part of nature is by definition smaller than the whole of nature, and different from other parts. As a result, no description of nature by a set can possibly yield the number of particles or the dimensionality of space-time. As long as we insist on using space-time or Hilbert spaces for the description of nature, we *cannot* understand the number of dimensions or the number of particles. Page 103

That is not too bad, as we know already that nature is *not* made of parts. We know that parts are only approximate concepts. In short, if nature were made of parts, it could not be a unity, or a 'one.' On the other hand, if nature is a unity, it cannot have parts.* Nature cannot be separable exactly. It cannot be made of particles.

To sum up, nature cannot be a set. Sets are lists of distinguishable elements. When general relativity and quantum theory are unified, nature shows no elements: *nature stops being a set at Planck scales*. This result clarifies a discussion we started earlier in relation to classical physics. There we discovered that matter objects were defined using space and time, and that space and time were defined using objects. Along with the results of quantum theory, this implies that in modern physics particles are defined in terms of the vacuum and the vacuum in terms of particles. Circularity is not a good idea, but we can live with it – at low energy. But at Planck energy, vacuum and particles are indistinguishable from each other. Particles and vacuum – thus everything – are the same. We have to abandon the circular definition. This is a satisfactory outcome; however, it also implies that nature is not a set.

Also space-time duality implies that space is not a set. Space-time duality implies that events cannot be distinguished from each other, and thus do not form elements of some space. Phil Gibbs has given the name *event symmetry* to this property of nature. Ref. 121 This thought-provoking term, although still containing the term 'event', emphasizes the impossibility to use a set to describe space-time.

In short,

* As a curiosity, practically the same discussion can already be found in Plato's *Parmenides*, written in the fourth century BCE. There, Plato musically ponders different arguments on whether nature is or can be a *unity* or a *multiplicity*, i.e., a set. It seems that the text is based on the real visit to Athens by Parmenides and Zeno. (Their home city, Elea, was near Naples.) Plato does not reach a conclusion. Modern physics, however, does. Ref. 120

▷ Nature cannot be made of vacuum and particles.

This is a bizarre result. Atomists, from Democritus to Galileo, have been persecuted throughout history. Were their battles all in vain? Let us continue to clarify our thoughts.

Does the universe contain anything?

To state that the universe contains something implies that we are able to distinguish the universe from its contents. However, we now know that precise distinctions are impossible. If nature is not made of parts, it is wrong to say that the universe *contains* something.

Let us go further. We need a description of nature that allows us to state that at Planck energy nothing can be distinguished from anything else. For example, it must be impossible to distinguish particles from each other or from the vacuum. There is only one solution: everything – or at least, what we call 'everything' in everyday life – must be made of the same single constituent. All particles are made of one 'piece'. Every point in space, every event, every particle and every instant of time must be made of the same single constituent.

An amoeba

> "A theory of everything describing nothing is not better than a theory of nothing describing everything."
>
> Anonymous

We have found that parts are approximate concepts. The parts of nature are not strictly smaller than nature itself. As a result, any 'part' must be extended. Let us try to extract some more information about the constituents of nature.

In any unified theory, all the concepts that appear must be only *approximately* parts of the whole. Thus we need an entity Ω, describing nature, which is not a set but which can be approximated by one. This is strange. We are all convinced very early in our lives that we are a *part* of nature. Our senses provide us with this information. We are not used to thinking otherwise. But now we have to.

Let us straight away eliminate a few options for Ω. One concept without parts is the empty set. Perhaps we need to construct a description of nature from the empty set? We could be inspired by the usual construction of the natural numbers from the empty set.

Vol. III, page 291

However, the empty set makes only sense as the opposite of some full set. So the empty set is not a candidate for Ω.

Another possible way to define approximate parts is to construct them from multiple copies of Ω. But in this way we would introduce a new set through the back door. Furthermore, new concepts defined in this way would not be approximate.

We need to be more imaginative. How can we describe a whole which has no parts, but which has parts approximately? Let us recapitulate. The world must be described by a single entity, sharing all properties of the world, but which can be approximated as a set of parts. For example, the approximation should yield a set of space points and a set of particles. But also, whenever we look at any 'part' of nature, without any approximation, we should not be able to distinguish it from the whole world. Composite objects are not always larger than their constituents. On the other hand, composed objects must usually *appear* to be larger than their constituents. For example, space 'points' or 'point' particles are tiny, even though they are only approximations. Which concept without boundaries can be at their origin? Using usual concepts, the world is everywhere at the same time; if nature is to be described by a single constituent, this entity must be extended.

The entity has to be a single one, but it must *seem* to be multiple: it has to be multiple approximately, as nature shows multiple aspects. The entity must be something folded. It must be possible to count the folds, but only approximately. (An analogy is the question of how many grooves there are on an LP or a CD: depending on the point of view, local or global, one gets different answers.) Counting folds would correspond to a length measurement.

The simplest model would be a single entity which is extended and fluctuating, reaches spatial infinity, allows approximate localization, and thus allows approximate definition of parts and points.* In more vivid imagery, nature could be described by some deformable, folded and tangled entity: a giant, knotted amoeba. An amoeba slides between the fingers whenever we try to grab a part of it. A perfect amoeba flows around any knife trying to cut it. The only way to hold it would be to grab it in its entirety. However, for someone himself made of amoeba strands, this is impossible. He can only grab it approximately, by catching part of it and approximately blocking it, for example using a small hole, so that the escape takes a long time.

Summary of the fourth argument for extension

The lack of particles and of sets in nature leads to describing nature by a single constituent. Nature is thus modelled by an entity which is *one single 'object'* (to eliminate distinguishability), which is *extended* (to eliminate localizability) and which is *fluctuating* (to ensure approximate continuity). Nature is a far-reaching, fluctuating fold. Nature is similar to an amoeba. The tangled branches of the amoeba allow a definition of length via counting of the folds. In this way, *discreteness* of space, time, and particles could also be realized; the quantization of space-time, matter and radiation thus follows. Any flexible and deformable entity is also a perfect candidate for the realization of diffeomorphism invariance, as required by general relativity.

A simple candidate for the extended fold is the image of a fluctuating, flexible *tube* of Planck diameter. Counting tubes implies determining distances or areas. The minimum possible count (one) gives the minimum distance, from which quantum theory is derived. In fact, at this point we can use as a model any flexible object with a small dimension, such as a tube, a thin sheet, a ball chain or a woven collection of rings. We will explore these options below.

Page 148

Challenge 95 r

* This is the simplest model; but is it the only way to describe nature?

THE ENTROPY OF BLACK HOLES

We are still collecting arguments to determining the shape of fundamental constituents. Another approach is to study situations where particles appear in large numbers. Systems composed of many particles behave differently depending on whether the particles are point-like or extended. In particular, their entropy is different. Studying large-number entropy thus allows us to determine the shape of components. The most revealing situations are those in which large numbers of particles are crammed in a small volume. Therefore we are led to study the entropy of black holes. Indeed, black holes tell us a lot about the fundamental constituents of nature.

A black hole is a body whose gravity is so strong that even light cannot escape. It is easily deduced from general relativity that any body whose mass m fits inside the so-called Schwarzschild radius

$$r_S = 2Gm/c^2 \tag{115}$$

is a black hole. A black hole can be formed when a whole star collapses under its own weight. Such a black hole is a macroscopic body, with a large number of constituents. Therefore it has an entropy. The entropy S of a macroscopic black hole was determined
Ref. 57, Ref. 58 by Bekenstein and Hawking, and is given by

$$S = \frac{k}{4l_{Pl}^2} A = \frac{kc^3}{4\hbar G} A \quad \text{or} \quad S = k\frac{4\pi G m^2}{\hbar c} \tag{116}$$

where k is the Boltzmann constant and $A = 4\pi r_S^2$ is the surface of the black hole horizon.
Ref. 122 This important result has been derived in many different ways. The various derivations confirm that space-time and matter are equivalent: they show that the entropy value can be interpreted as an entropy either of matter or of space-time. In the present context, the two main points of interest are that the entropy is *finite*, and that it is *proportional to the area* of the black hole horizon.

In view of the existence of minimum lengths and times, the finiteness of the entropy is not surprising: it confirms the idea that matter is made of a finite number of discrete constituents per given volume (or area). It also shows that these constituents behave statistically: they fluctuate. In fact, quantum gravity implies a finite entropy for any object,
Ref. 35 not only for black holes. Jacob Bekenstein has shown that the entropy of an object is always smaller than the entropy of a (certain type of) black hole of the same mass.

The entropy of a black hole is proportional to its horizon area. Why? This question
Ref. 123 has been the topic of a stream of publications.* A simple way to understand the entropy–surface proportionality is to look for other systems in nature whose entropy is proportional to system surface instead of system volume. In general, the entropy of a collection
Ref. 125 of flexible one-dimensional objects, such as polymer chains, shares this property. Indeed, the entropy of a polymer chain made of N monomers, each of length a, whose ends are

Ref. 124 * The result can be derived from quantum statistics alone. However, this derivation does not yield the proportionality coefficient.

Ref. 126 kept a distance r apart, is given by

$$S(r) = k\frac{3r^2}{2Na^2} \quad \text{for} \quad Na \gg \sqrt{Na} \gg r \,. \tag{117}$$

This formula can be derived in a few lines from the properties of a random walk on a lattice, using only two assumptions: the chains are extended; and they have a characteristic internal length a given by the smallest straight segment. Expression (117) is only valid if the polymers are effectively infinite: in other words, if the length Na of the chain and the *elongation* $a\sqrt{N}$, are much larger than the radius r of the region of interest. If the chain length is comparable to or smaller than the region of interest, we get the usual extensive entropy, satisfying $S \sim r^3$. Thus *only flexible extended constituents yield an* $S \sim r^2$ *dependence.*

However, there is a difficulty. From the expression for the entropy of a black hole we deduce that the elongation $a\sqrt{N}$ is given by $a\sqrt{N} \approx l_{\text{Pl}}$; thus it is much smaller than the radius of a general macroscopic black hole, which can have a diameter of several kilometres. On the other hand, the formula for long constituents is only valid when the chains are longer than the distance r between the end points.

This difficulty disappears when we remember that space near a black hole is strongly curved. All lengths have to be measured in the same coordinate system. It is well known that for an outside observer, any object of finite size falling into a black hole seems to cover the complete horizon for long times (whereas for an observer attached to the object Ref. 109 it falls into the hole in its original size). In short, an extended constituent can have a proper length of Planck size but still, when seen by an outside observer, be as long as the horizon of the black hole. We thus find

▷ Black holes are made of extended constituents.

Another viewpoint can confirm this result. Entropy is (proportional to) the number of yes-or-no questions needed to know the exact state of the system. But if a system is defined by its surface, as a black hole is, its components must be extended.

Finally, imagining black holes as made of extended constituents is also consistent with the so-called *no-hair theorem*: black holes' properties do not depend on what falls into them – as long as all matter and radiation particles are made of the same extended components. The final state of a black hole only depends on the number of extended constituents.

Summary of the fifth argument for extension

Black hole entropy is best understood as resulting from extended constituents that tangle Page 289 and fluctuate. And black hole entropy confirms that vacuum and particles are made of common constituents.

EXCHANGING SPACE POINTS OR PARTICLES AT PLANCK SCALES

Let us now focus on the exchange behaviour of fundamental constituents in nature. We saw above that 'points' in space have to be abandoned in favour of continuous, fluctuating constituents common to space, time and matter. Is such a constituent a boson or a fermion? If we exchange two points of empty space, in everyday life, nothing happens. Indeed, at the basis of quantum field theory is the relation

$$[x, y] = xy - yx = 0 \tag{118}$$

between any two points with coordinates x and y, making them bosons. But at Planck scales, because of the existence of minimal distances and areas, this relation must at least be changed to

$$[x, y] = l_{\mathrm{Pl}}^2 + \ldots \,. \tag{119}$$

This means that 'points' are neither bosons nor fermions.* 'Points' have more complex exchange properties. In fact, the term on the right-hand side will be energy-dependent, to an increasing extent as we approach Planck scales. In particular, as we have seen, Ref. 47 gravity implies that a double exchange does not lead back to the original situation at Planck scales.

Constituents obeying this or similar relations have been studied in mathematics for Ref. 127 many decades: they are called *braids*. Thus space is not made of points at Planck scales, but of braids or their generalizations, namely tangles. We find again that quantum theory and general relativity taken together imply that the vacuum must be made of extended constituents.

We now turn to particles. All particles in nature behave in a similar way: we know that at low, everyday energies, particles of the same type are *identical*. Experiments sensitive to quantum effects show that there is no way to distinguish them: any system of several identical particles has permutation symmetry. On the other hand, we know that at Planck energy all low-energy symmetries disappear. We also know that at Planck energy permutation cannot be carried out, as it implies exchanging positions of two particles. At Planck energy, nothing can be distinguished from vacuum; thus no two entities can be shown to have identical properties. Indeed, no two particles can be shown to be indistinguishable, as they cannot even be shown to be separate.

What happens when we slowly approach Planck energy? At everyday energies, permutation symmetry is defined by commutation or anticommutation relations between particle creation operators

$$a^{\dagger} b^{\dagger} \pm b^{\dagger} a^{\dagger} = 0 \,. \tag{120}$$

At Planck energy this cannot be correct. Quantum gravity effects modify the right-hand side: they add an energy-dependent term that is negligible at experimentally accessible energies but becomes important at Planck energy. We know from our experience with

* The same reasoning applies to the so-called fermionic or Grassmann coordinates used in supersymmetry. They cannot exist at Planck energy.

Ref. 47 Planck scales that, in contrast to everyday life, exchanging particles twice cannot lead back to the original situation. A double exchange at Planck energy cannot have no effect, because at Planck energy such statements are impossible. The simplest extension of the commutation relation (120) for which the right-hand side does not vanish is *braid symmetry*. This again suggests that particles are made of extended constituents. Ref. 127

Summary of the sixth argument for extension

Extrapolating both point and particle indistinguishability to Planck scales suggests extended, braided or tangled constituents.

The meaning of spin

As last argument we will now show that the extension of particles makes sense even at everyday energy. Any particle is a part of the universe. A part is something that is different from anything else. Being 'different' means that exchange has some effect. *Distinction means detection of exchange.* In other words, any part of the universe is also described by its *exchange behaviour*.

In nature, particle exchange is composed of rotations. In other words, parts of nature are described by their *rotation behaviour*. This is why, for microscopic particles, exchange behaviour is specified by spin. *Spin distinguishes particles from vacuum.**

We note that volume does not distinguish vacuum from particles; neither does rest mass or charge: nature provides particles without measurable volume, rest mass or charge, such as photons. The only observables that distinguish particles from vacuum are spin and momentum. In fact, linear momentum is only a limiting case of angular momentum. We thus find again that rotation behaviour is the basic aspect distinguishing particles from vacuum.

If spin is the central property that distinguishes particles from vacuum, finding a model for spin is of central importance. But we do not have to search for long. A model for spin 1/2 is part of physics folklore since almost a century. Any belt provides an example, as we discussed in detail when exploring permutation symmetry. Vol. IV, page 112 Any localized structure with any number of tails or tethers attached to it – tails or tethers that reach the border of the region of space under consideration – has the same properties as a spin 1/2 particle. The only condition is that the tails or tethers themselves are *unobservable*. It is a famous exercise to show that such a model, like one of those shown in Figure 9, is indeed invariant under 4π rotations but not under 2π rotations, and that two such particles get entangled when exchanged, but get untangled when exchanged twice. Such a *tether model* has all the properties of spin 1/2 particles, independently of the precise structure of the central region, which is not important at this point. The tether model even has the same problems with highly curved space as real spin 1/2 particles have. We will explore the issues in more detail shortly. Page 174

* With a flat (or other) background, it is possible to define a *local* energy–momentum tensor. Thus particles can be defined. Without a background, this is not possible, and only global quantities can be defined. Without a background, even particles cannot be defined. Therefore, in this section we assume that we have a slowly varying space-time background.

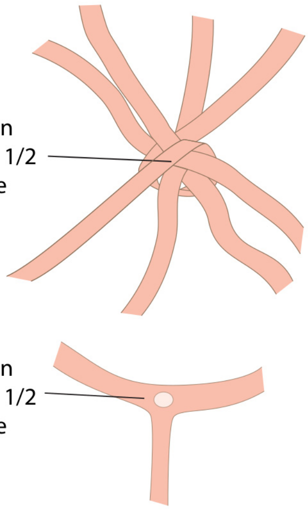


FIGURE 9 Possible models for a spin 1/2 particle.

The tether model thus confirms that rotation is partial exchange. More interestingly, Ref. 128 it shows that rotation implies connection with the border of space. Extended particles can be rotating. Particles can have spin 1/2 provided that they have tethers going to the border of space. If the tethers do not reach the border, the model does not work. Spin 1/2 thus even seems to *require* extension.

Challenge 96 e It is not hard to extend this idea to include spin 1 particles. In short, both bosons and fermions can be modelled with extended constituents.

Summary of the seventh argument for extension

Exploring the properties of particle spin suggests the existence of extended, unobservable constituents in elementary fermions. We note that gravitation is not used explicitly in the argument. It is used implicitly, however, in the definition of the locally flat space-time and of the asymptotic region to where the tethers are reaching.

CURIOSITIES AND FUN CHALLENGES ABOUT EXTENSION

> No problem is so big or complicated that it can't be run away from.
>
> Charles Schulz

In case that this section has not provided enough food for thought, here is some more.

* *

Quantum theory implies that even if tight walls exist, the lid of a box made of them could Challenge 97 s never be tightly shut. Can you provide the argument?

* *

Challenge 98 e Can you provide an argument against the idea of extended constituents in nature? If so, publish it!

* *

Does duality imply that the cosmic background fluctuations (at the origin of galaxies and clusters) are the same as vacuum fluctuations? Challenge 99 s

* *

Does duality imply that a system with two small masses colliding is equivalent to a system with two large masses gravitating? Challenge 100 s

* *

It seems that in all arguments so far we have assumed that time is continuous, even though we know it is not. Does this change the conclusions? Challenge 101 d

* *

Duality also implies that in some sense large and small masses are equivalent. A mass m in a radius r is equivalent to a mass $m_{\rm Pl}^2/m$ in a radius $l_{\rm Pl}^2/r$. In other words, duality transforms mass density from ρ to $\rho_{\rm Pl}^2/\rho$. Vacuum and maximum density are equivalent! Vacuum is thus dual to black holes.

* *

Total symmetry and space-time duality together imply that there is a symmetry between all values an observable can take. Do you agree? Challenge 102 s

* *

Any description is a mapping from nature to mathematics, i.e., from observed differences (and relations) to thought differences (and relations). How can we do this accurately, if differences are only approximate? Is this the end of physics? Challenge 103 s

* *

Duality implies that the notion of initial conditions for the big bang makes no sense, as we saw earlier by considering the minimal distance. As duality implies a symmetry between large and small energies, the big bang itself becomes a vague concept. What else do extended constituents imply for the big bang? Challenge 104 d

* *

Can you show that going to high energies or selecting a Planck-size region of space-time is equivalent to visiting the big bang? Challenge 105 d

* *

In 2002, Andrea Gregori made a startling prediction for any model using extended constituents that reach the border of the universe: if particles are extended in this way, their mass should depend on the size of the universe. Thus particle masses should change with time, especially around the big bang. Is this conclusion unavoidable? Ref. 129 Challenge 106 s

* *

What is wrong with the following argument? We need lines to determine areas, and we need areas to determine lines. This implies that at Planck scales, we cannot distinguish areas from lengths at Planck scales.

Challenge 107 s

* *

Ref. 129, Ref. 130
Challenge 108 s

We need a description for the expansion of the universe in terms of extended constituents. Various approaches are being explored. Can you speculate about the solution?

Gender preferences in physics

Why has extension appeared so late in the history of physics? Here is a *not too serious* answer. When we discussed the description of nature as made of tiny balls moving in a void, we called this as a typically male idea. This implies that the female part is missing. Which part would that be?

Vol. I, page 339

From a general point of view, the female part of physics might be the quantum description of the vacuum, the container of all things. We can speculate that if women had developed physics, the order of its discoveries might have been different. Instead of studying matter first, as men did, women might have studied the vacuum first. And women might not have needed 2500 years to understand that nature is not made of a void and little balls, but that everything in nature is made of extended constituents. It is curious that (male) physics took so long for this discovery.

CHECKS OF EXTENSION

The idea that nature is described by extended constituents is taken for granted in all current research approaches to unification. How can we be sure that extension is correct? The arguments presented so far provide several possible checks. We start with some options for *theoretical* falsification.

— Any explanation of black hole entropy *without* extended constituents would invalidate the need for extended constituents.
— A single thought experiment *invalidating* extended constituents would prove extension wrong.
— Extended constituents must appear if we start from *any* physical (low-energy) concept – not only from length measurements – and study how the concept behaves at Planck scales.
— Invalidating the requirement of extremal identity, or duality, would invalidate the need for extended constituents. As Edward Witten likes to say, any unified model of nature must include duality.
— If the measurement of length could be shown to be *unrelated* to the counting of folds of extended constituents, extension would become unnecessary.
— Finding any property of nature that *contradicts* extended constituents would spell the end of extension.

Challenge 109 e
Page 113

Any of these options would signal the end for almost all current unification attempts.

Fortunately, theoretical falsification has not yet occurred. But physics is an experimental science. What kind of *data* could falsify the idea of extended constituents?

— Observing a single particle in cosmic rays with energy above the corrected Planck energy would invalidate the invariant limits and thus also extension. However, the present particle energy record, about 0.35 ZeV, is a million times lower than the Planck energy. Ref. 131
— Paul Mende has proposed a number of checks on the motion of extended objects in space-time. He argues that an extended object and a mass point move differently; the differences could be noticeable in scattering or dispersion of light near masses. Ref. 132

Experimental falsification of extension has not yet occurred. In fact, experimental falsification is rather difficult. It seems easier and more productive to *confirm* extension. Confirmation is a well-defined project: it implies to deduce all those aspects of nature that are given in the millennium list of unexplained properties. Among others, confirmation requires to find a concrete model, based on extended constituents, for the electron, the muon, the tau, the neutrinos, the quarks and all bosons. Confirmation also requires using extended constituents to realize an old dream of particle physics: to deduce the values of the coupling constants and particle masses. Before we attempt this deduction, we have a look at some other attempts. Page 19

Current research based on extended constituents

> “To understand is to perceive patterns.”
>
> Isaiah Berlin*

The Greeks deduced the existence of atoms from the observation that fish can swim through water. They argued that only if water is made of atoms could a fish make its way through it, by pushing the atoms aside. We can ask a similar question of a particle flying through a vacuum: why is it able to do so? A vacuum cannot be a fluid or a solid composed of small constituents, as its dimensionality would not then be fixed. Only one possibility remains: both vacuum and particles are made of extended constituents. Ref. 114

The idea of describing matter as composed of extended constituents dates from the 1960s. That of describing nature as composed of ‘infinitely’ extended constituents dates from the 1980s. In addition to the arguments presented so far, current research provides several other approaches that arrive at the same conclusion.

* *

Bosonization, the construction of fermions using an infinite number of bosons, is a central aspect of modern unification attempts. It also implies coupling duality, and thus the extension of fundamental constituents. Ref. 133

* *

Research into quantum gravity – in particular the study of spin networks, spin foams and loop quantum gravity – has shown that the vacuum can be thought of as a collection of extended constituents. Ref. 134, Ref. 135

* Isaiah Berlin (b. 1909 Riga, d. 1997 Oxford) was an influential political philosopher and historian of ideas.

* *

Ref. 136 In the 1990s, Dirk Kreimer showed that high-order QED Feynman diagrams are related to knot theory. He thus proved that extension arrives by the back door even when electromagnetism is described in terms of point particles.

* *

Ref. 137 A popular topic in particle physics, 'holography', relates the surface and the volume of physical systems at high energy. It implies extended constituents of nature.

* *

Vol. IV, page 157 It is long known that wave function collapse can be seen as the result of extended constituents. We will explore the details below.

* *

Ref. 138, Ref. 139 Ref. 140, Ref. 141 Ref. 142, Ref. 143 Page 347 At the start of the twenty-first century, a number of new approaches to describe elementary particles appeared, such as models based on string nets, models based on bands, models based on ribbons, and models based on knots. All these attempts make use of extended constituents. Several of them are discussed in more detail below.

Despite the use of extension, none of these attempts solved a single problem from the millennium list. One approach – especially popular between the years 1984 and 2010 – merits a closer look.

Superstrings – extension plus a web of dualities

> "Throw physic to the dogs; I'll none of it."
> William Shakespeare, *Macbeth*.

Ref. 144 *Superstrings* and *supermembranes* – often simply called *strings* and *membranes* – are extended constituents in the most investigated physics conjecture ever. The approach contains a maximum speed, a minimum action and a maximum force (or tension). The approach thus incorporates special relativity, quantum theory and general relativity. This attempt to achieve the unified description of nature uses four ideas that go beyond standard general relativity and quantum theory:

1. Particles are conjectured to be *extended*. Originally, particles were conjectured to be one-dimensional *oscillating superstrings*. In a subsequent generalization, particles are conjectured to be fluctuating higher-dimensional *supermembranes*.
2. The conjecture uses *higher dimensions* to unify interactions. A number of space-time dimensions much higher than 3 + 1, typically 10 or 11, is necessary for a mathematically consistent description of superstrings and membranes.
3. The conjecture is based on *supersymmetry*. Supersymmetry is a symmetry that relates matter to radiation, or equivalently, fermions to bosons. Supersymmetry is the most general local interaction symmetry that can be constructed mathematically. Supersymmetry is the reason for the terms 'superstring' and 'supermembrane'.
4. Ref. 145 The conjecture makes heavy use of *dualities*. In the context of high-energy physics, dualities are symmetries between large and small values of physical observables. Important examples are space-time duality and coupling constant duality. Dualities are

global interaction and space-time symmetries. They are essential for the inclusion of gauge interaction and gravitation in the quantum description of nature. Dualities also express a fundamental equivalence between space-time and matter–radiation. Dualities also imply and contain *holography*, the idea that physical systems are completely fixed by the states on their bounding surface.

By incorporating these four ideas, the *superstring conjecture* – named so by Brian Greene, one of its most important researchers – acquires a number of appealing characteristics. Ref. 146

Why superstrings and supermembranes are so appealing

First of all, the superstring conjecture is unique: the Lagrangian is claimed to be unique and to have no adjustable parameters. Furthermore, as we would expect from a description involving extended constituents, the conjecture includes gravity. In addition, the conjecture describes interactions: it describes gauge fields. The conjecture thus expands quantum field theory, while retaining all its essential points. In this way, the conjecture fulfils most of the requirements for a unified description of motion that we have deduced so far. For example, particles are not point-like, there are minimal length and time intervals, and all other limit quantities appear. (However, sets are still used.)

The superstring conjecture has many large symmetries, which arise from the many dualities it contains. These symmetries connect many situations that seem intuitively to be radically different: this makes the conjecture extremely fascinating, but also difficult to picture.

The conjecture shows special cancellations of anomalies and of other inconsistencies. Historically, the first example was the Green–Schwarz anomaly cancellation; superstrings also solve other anomalies and certain inconsistencies of quantum field theory.

Edward Witten, the central figure of the field, liked to say that quantum theory cures the infinities that appear in e^2/r when the distance r goes to zero; in the same way, superstrings cure the infinities that appear in m^2/r when the distance r goes to zero.*

Also following Witten, in the superstring conjecture, the interactions follow from the particle definitions: interactions do not have to be added. That is why the superstring conjecture predicts gravity, gauge theory, supersymmetry and supergravity.

About gravity, one of the pretty results of the superstring conjecture is that super-
Ref. 147 strings and black holes are complementary to each other. This was argued by Polchinsky, Horowitz and Susskind. As expected, superstrings explain the entropy of black holes.
Ref. 122 Strominger and Vafa showed this in 1996.

The superstring conjecture naturally includes *holography*, the idea that the degrees of freedom of a physical system are determined by its boundary. In particular, holography provides for a deep duality between gauge theory and gravity. More precisely, there is a correspondence between quantum field theory in flat space and the superstring conjecture in certain higher-dimensional spaces that contain anti-de Sitter space.

In short, the superstring conjecture implies fascinating mathematics. Conformal invariance enters the Lagrangian. Concepts such as the Virasoro algebra, conformal field theory, topological field theory and many related ideas provide vast and fascinating generalizations of quantum field theory.

* This argument is questionable, because general relativity already cures that divergence.

Why the mathematics of superstrings is difficult

The superstring conjecture, like all modern descriptions of physics, is claimed to be described by a Lagrangian. The Lagrangian is constructed starting from the Lagrangian for the motion of a classical superstring of matter. Then the Lagrangian for the corresponding quantum superstring fields is constructed, and then higher dimensions, supersymmetry, dualities and membranes are incorporated. This formulation of the superstring conjecture takes for granted the existence of a space-time background.

The Lagrangian of the superstring conjecture is extremely complex, much too complex to write it down here. It is not as simple as the Lagrangian of the standard model of particle physics or the Lagrangian of general relativity. But the complexity of the Lagrangian is not the only reason why the studying the superstring conjecture is difficult.

It turns out that exploring how the known 4 dimensions of space-time are embedded in the 10 or 11 dimensions of the superstring conjecture is extremely involved. The topology and the size of the additional dimensions is unclear. There are only few people who are able to study these options.

Indeed, a few years ago a physicist and a mathematician listened to a talk on superstrings, describing nature in eleven dimensions. The mathematician listened intensely and obviously enjoyed the talk. The physicist did not understand anything and got more and more annoyed. At the end, the physicist had a terrible headache, whereas the mathematician was full of praise. 'But how can you even understand this stuff?', asked the physicist. 'I simply picture it in my head!' 'But *how* do you imagine things in eleven dimensions?' 'Easy! I first imagine them in N dimensions and then let N go to 11.'

Testing superstrings: couplings and masses

Ref. 148 One of the main results of quantum chromodynamics or QCD, the theory of strong interactions, is the explanation of mass relations such as

$$m_{\text{proton}} \sim \mathrm{e}^{-k/\alpha_{\text{Pl}}}\, m_{\text{Pl}} \quad \text{and} \quad k = 11/2\pi\,,\ \alpha_{\text{Pl}} \approx 1/25\,. \tag{121}$$

Here, the value of the strong coupling constant α_{Pl} is taken at the Planck energy. In other words, a general understanding of masses of bound states of the strong interaction, such as the proton, requires little more than a knowledge of the Planck energy and the coupling constant at that energy. The approximate value $\alpha_{\text{Pl}} \approx 1/25$ is an empirical value based on experimental data. Page 382

Any unified theory must allow us to calculate the three gauge coupling constants as a function of energy, thus also α_{Pl}. In the superstring conjecture, this calculation faces two issues. First, the coupling constants must be the same for all particles; equivalently, electric, weak and strong charge must be quantized. It is not clear whether superstrings fullfil this condition. There does not seem to be any paper addressing the issue.

Secondly, in order to calculate coupling constants and particle masses, the vacuum state must be known. However, it is not. The search for the vacuum state – the precise embedding of four dimensions in the total ten – is the main difficulty facing the superstring conjecture.

The vacuum state of the superstring conjecture is expected to be one of an rather

Ref. 149 involved set of topologically distinct manifolds. It was first estimated that there are only 10^{500} possible vacuum states; recent estimates raised the number to $10^{272\,000}$ candidate vacuum states. Since the universe contains 10^{80} atoms, it seems easier to find a particular atom somewhere in the universe than to find the correct vacuum state. The advantages due to a unique Lagrangian are thus lost.

The calculation of particle masses faces a further issue. The superstring conjecture predicts states with Planck mass and with zero mass. The zero-mass particles are then thought to get their actual mass, which is tiny compared with the Planck mass, from the Higgs mechanism. However, the Higgs mechanism and its measured properties – or any other parameter of the standard model – have not yet been deduced from superstrings.

The status of the superstring conjecture

> Es ist nichts Großes ohne Leidenschaft vollbracht worden, noch kann es ohne solche vollbracht werden.*
>
> Friedrich Hegel, *Enzyklopädie.*

After several decades of research, superstring researchers are stuck. Despite the huge collective effort by extremely smart researchers, not a single calculation of an *experimentally measurable* value has been performed. For example, the superstring conjecture has not predicted the masses of any elementary particle, nor the value of any coupling constant, nor the number of gauge interactions, nor the number of particle generations. So far, Page 19 *none* of the open issues in physics that are listed the millennium list has been solved by the superstring conjecture. This disappointing situation is the reason that many scholars, Ref. 150 including several Nobel Prize winners, dismiss the superstring conjecture altogether.

What are the reasons that the superstring conjecture, like several other approaches based on extended constituents, was unsuccessful? First of all, superstrings and supermembranes are *complex*: superstrings and supermembranes move in many dimensions, carry mass, have tension and carry (supersymmetric) fields. In fact, the precise mathematical definition of a superstring or a supermembrane and their features is so complex that already understanding the definition is beyond the capabilities of most physicists. But a high complexity always nourishes the doubt that some of the underlying assumptions do not apply to nature. For example, it is difficult to imagine that a macroscopic concept like 'tension' applies at the fundamental level: in nature, tension only applies to systems that have parts, i.e., to composed systems; tension cannot be a property of fundamental entities.

Superstrings are complex entities. And no researcher tried to make them simple. Put in different terms, the superstring conjecture was not successful because its *basic prin-* Ref. 151 *ciples* have never been clarified. It is estimated that, from 1984 to 2010, over 10 000 man-years have been invested in the exploration of the superstring conjecture. Compare this with about a dozen man-years for the foundations and principles of electrodynamics, a dozen man-years for the foundations and principles of general relativity, and a dozen man-years for the foundation and principles of quantum theory. The lack of clear foundations of the superstring conjecture is regularly underlined even by its supporters, such

* 'Nothing great has been achieved without passion, nor can it be achieved without it.' Hegel, an influential philospher, writes this towards the end of the third and last part of his *Enzyklopädie der philosophischen Wissenschaften im Grundrisse*, §474, 296.

as Murray Gell-Mann. And despite this gap, *no* research papers on the basic principles exist – to this day.

Apart from the complexity of the conjecture, a further aspect about superstrings and supermembranes has been getting growing attention: the original claim that there is a unique well-defined Lagrangian has been retracted; it not even made by the most outgoing proponents any more. In other words, it is not clear which specific supermembrane conjecture should be tested against experiment in the first place.

These developments effectively dried out the research field. At the latest since 2014 *Strings* conference, it became clear that the string research community has quietly given up its quest to achieve a unified theory with the help of superstrings or supermembranes. Several prominent researchers are now looking for other microscopic models of nature.

SUMMARY ON EXTENSION IN NATURE

> Wir müssen wissen, wir werden wissen.*
> David Hilbert

We have explored nature at her limits: we have studied the Planck limits, explored three-dimensionality, curvature, particle shape, renormalization, spin and bosonization; we have investigated the cosmological constant problem and searched for a 'background-
Ref. 152
free' description of nature. As a result, we have found that at Planck scales, all these explorations lead to the same conclusions:

— Matter and vacuum are two sides of the same medal.
— Points and sets do not describe nature correctly.
— What we usually call space-time points and point particles are in fact made up of common and, above all, *extended* constituents.

We can reach the conclusions in an even simpler way. What do quantum theory and black holes have in common? They both suggest that nature is made of extended entities. We will confirm shortly that both the Dirac equation and black hole entropy imply that particles, space and horizons are built from extended constituents.

Despite using extension as fundamental aspect, and despite many interesting results, *all* the attempts from the twentieth century – including the superstring conjecture, most quantum gravity models, supersymmetry and supergravity – have *not been successful* in understanding or in describing nature at the Planck scale. The reasons for this lack of success were the unclear relation to the Planck scale, the lack of clear principles, the use of incorrect assumptions, the use of sets and, above all, the unclear connection to experiment.

To be successful, we need a different approach to calculations with extended constituents. We need an approach that is built on Planck units, is based on clear principles, has few but correct assumptions, and provides predictions that stand up against experimental tests.

In our quest for a complete, unified theory of physics, one way to advance is by raising the following issue. The basis for the superstring conjecture is formed by four assump-

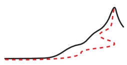
* 'We must know, we will know.' This was Hilbert's famous personal credo.

tions: extension, duality, higher dimensions and supersymmetry. Can we dispense with any of them? Now, duality is closely related to extension, for which enough theoretical and experimental evidence exists, as we have argued above. On the other hand, the expressions for the Schwarzschild radius and for the Compton wavelength imply, as we Page 72, page 78 found out earlier on, that the dimensionality of space and the statistics of particles are *undefined* at Planck scales. In other words, nature does not have higher dimensions nor Page 75, page 78 supersymmetry at Planck scales. Indeed, all experiments so far confirm this conclusion. In our quest for a complete theory of motion, we therefore drop the two incorrect assumptions and continue our adventure.

In summary, *extension* is the central property of the fundamental entities of nature that make up space, horizons, particles and interactions at Planck scales. We can thus phrase our remaining quest in the following specific way:

▷ How do extended entities relate the Planck constants c, $\hbar$, k and G to the electromagnetic, the weak and the strong interactions?

This question is rarely asked so specifically. Attempts to answer it are even rarer. (Can Challenge 110 e you find one?) Up to this point, we discovered: *Finding the Planck origin of the gauge interactions using extension means finding the complete, unified theory.* To be successful in this quest, we need three resources: simplicity, playfulness and intrepidity.



Chapter 7

THE BASIS OF THE STRAND CONJECTURE

> “We haven’t got the money, so we have to think.”
> Ernest Rutherford**

The two extremely precise descriptions of motion that were discovered in the twentieth century – quantum field theory and general relativity – are the low-energy approximations of how nature behaves at Planck scales. In order to understand nature at Planck scales, and thus to find the unified and complete description of motion, we follow the method that has been the most effective during the history of physics: we search for the *simplest* possible description. Simplicity was used successfully, for example, in the discovery of special relativity, in the discovery of quantum theory, and in the discovery of general relativity. We therefore use the guidance provided by simplicity to deduce a promising speculation for the unified and final theory of motion.

Requirements for a unified theory

The central requirement for any *unified* description is that it leads from Planck scales, and thus from Planck units, to quantum field theory, to the standard model of elementary particles and to general relativity. Page 164 In simple terms, as detailed below, the unified description must be valid for *all observations* and provide *complete precision*.

From the preceding chapters, we know already quite a bit about the unified description. In particular, any unified description of general relativity and quantum theory must use *extended* constituents. We discovered a number of reasons that are central for this conclusion. All these reasons appear only when quantum theory and general relativity are combined. First of all, only constituents that are extended allow us to deduce black hole entropy. Secondly, only extended constituents allow us to model that elementary particles are not point-like or that physical space is not made of points. Thirdly, only extended constituents allow us to model a smallest measurable space and time interval. Fourthly, only extended constituents allow us to model spin 1/2 in locally flat space-time.

But we are not only looking for a unified theory; we are also looking for the *complete* theory. This implies a second requirement: the theory must be *unmodifiable*. As we Page 166 will show below, if a candidate for a unified theory can be modified, or generalized, or reduced to special cases, or varied in any other way, it is not complete.

** Ernest Rutherford (b. 1871 Brightwater, d. 1937 Cambridge) was an important physicist and researcher who won the Nobel Prize in Chemistry for his work on atoms and radioactivity.

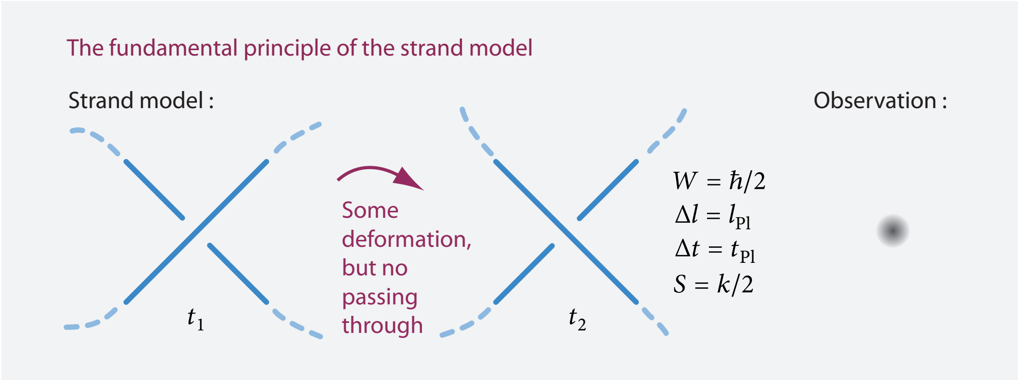


FIGURE 10 The fundamental principle of the strand conjecture: the simplest observation in nature, a ''point-like'' fundamental event, is defined by a crossing switch in three spatial dimensions. The crossing switch defines the action $\hbar/2$, the Planck length, the Planck time and half the Boltzmann constant $k/2$.

In the preceding chapters we have deduced many additional requirements that a complete theory must realize. The full list of requirements is given in Table 6. Certain requirements follow from the property that the description must be complete, others from the property that it must be unified, and still others from the property that it must describe nature with quantum theory and general relativity. More specifically, every requirement appears when the expressions for the Compton wavelength and for the Schwarzschild radius are combined. So far, the table is not found elsewhere in the research literature.

Challenge 111 e

TABLE 6 General requirements for a complete description of nature and of motion – as deduced so far.

Aspect	Requirements for the complete and unified description
Precision	must be *complete*; the unified description must precisely describe all motion – everyday, quantum and relativistic – and explain all open issues from the millennium list, given (again) in Table 8 on page 164, including the fine structure constant.
Modification	must be *impossible*, as explained on page 166.
Fundamental principles	must be *clear*. Otherwise the unified description is not falsifiable.
Vacuum and particles	*must not differ* at Planck scales, because of limits of measurement precision, as explained on page 65. Vacuum and particles therefore must be described by *common* fundamental constituents.
Fundamental constituents	must determine all observables.
Fundamental constituents	must be as *simple* as possible, to satisfy Occam's razor.
Fundamental constituents	must be *extended* and *fluctuating*, to explain black hole entropy, spin, minimum measurement intervals, space-time homogeneity and isotropy of space. See page 75.
Fundamental constituents	must be the *only unobservable* entities. If they were observable, the theory would not be complete, because the properties of the entities would need explanation. If additional unobservable entities would exist, the theory would be fiction, not science.

TABLE 6 (Continued) General requirements for a complete description of nature and of motion.

Aspect	Requirements for the complete and unified description
Physical points and sets	must *not exist* at Planck scales, as shown on page 67, page 106 and page 112 due to limits of measurement precision. Points and sets must only exist approximately, at everyday scales.
Evolution equations	must *not exist* at Planck scales, due to the lack of points and sets.
Non-locality	must be part of the description. Non-locality must be negligible at everyday scales, but important at Planck scales.
Physical systems	must not exist at Planck scales, due to limits of measurement precision. Systems must only exist approximately at everyday scales.
Universe	must not be a system, due to limits of measurement precision.
Big bang	must not be an event, and thus not be a beginning, as this would contradict the non-existence of points and sets in nature.
Singularities	must not exist, due to the limits of measurements.
Planck's natural units	must be *limit values* for each observable, within a factor of order one. Infinitely large or small measurement values must not exist.
Planck scale description	must imply quantum field theory, the standard model of particle physics, general relativity and cosmology.
Quantum field theory, including QED, QAD, QCD	must follow from the complete unified theory by eliminating G.
General relativity	must follow from the complete unified theory by eliminating $\hbar$.
Planck's natural units	must define all observables, including coupling constants.
Relation to experiment	must be as simple as possible, to satisfy Occam's razor.
Background dependence	is required, as background independence is logically impossible.
Background space-time	must be *equal* to physical space-time at everyday scale, but must *differ* globally and at Planck scales.
Circularity of definitions	of physical concepts must be part of the complete, unified description, as a consequence of being 'precise talk about nature'.
Axiomatic description	must be impossible, as nature is not described by sets. Hilbert's sixth problem must have no solution.
Dimensionality of space	must be *undefined* at Planck scales, as space is undefined there.
Symmetries	must be *undefined* at Planck scales, due to the limits to measurement precision.
Large and small scales	must be *similar*, due to the limits to measurement precision.

The detailed requirement list given in Table 6 can be considerably shortened. Shortening the list is possible because the various requirements are *consistent* with each other. In fact, shortening is possible because a detailed check confirms a suspicion that arose during the last chapters: extension alone is sufficient to explain all those requirements that seem particularly surprising or unusual, such as the lack of points or the lack of axioms. Such a shortened list also satisfies our drive for simplicity. After shortening, two

Challenge 112 e

requirements for a unified theory remain:

▷ *The complete theory must describe nature* at and below the Planck scale* *as made of extended constituents fluctuating in a background.* Extended constituents must explain particles, space, interactions and horizons.

▷ In the complete theory, the *fluctuations* of the extended constituents *must explain all motion.* The Planck-scale fluctuations must describe all observed examples of everyday, quantum and relativistic motion with complete precision, imply all interactions, all particles, all concepts of physics and explain all fundamental constants.

This requirement summary is the result of our journey up to this point. The summary forms the starting point for the final leg of our adventure. If you do not agree with these
Challenge 113 s two requirements, take a rest and explore your disagreement in all its details.

Looking at the table of requirements for the complete theory – both the full one and the shortened one – we note something astonishing. Even though all requirements appear when quantum physics and general relativity are combined, each of these requirements *contradicts* both quantum physics and general relativity. The final theory thus *differs* from both pillars of modern physics. A final theory cannot be found if we remain prisoners of either quantum theory or general relativity. To put it bluntly, each requirement for the final theory contradicts every result of twentieth century physics! This unexpected conclusion is the main reason that past attempts failed to discover the final theory. In fact, most past attempts do not fulfil the requirements because various scholars explicitly disagree with one or several of them.

The requirement of the *extension* of the fundamental constituents is the central result of the quest so far. A complete theory must make a statement about these constituents. The fundamental constituents, sometimes also called *fundamental degrees of freedom*, must explain everything we observe and know about nature. In particular, the constituents must explain the curvature of space, the entropy of black holes, the origin of gauge interactions and the spectrum, the masses and the other properties of all elementary particles. The fundamental constituents must be extended. Extension is the reason that the complete theory contradicts both general relativity and quantum theory; but extension must also yield these theories as excellent approximations. In short, extension is the key to finding the complete theory.

The requirement of fluctuating extended constituents resulted from our drive for extreme simplicity. Using this requirement, the search for a candidate final theory does
Ref. 153 not take long. Of the few candidates that satisfy the requirement, it appears that the
Page 167 *simplest* is the one based on *fluctuating featureless strands.* In this approach, *strands*,**

* The complete theory *must not* describe nature *beyond* the Planck scales, because that domain is not accessible by experiment. A more relaxed requirement is that the predictions of the theory must be *independent* of any fantasies of what one might imagine *beyond* Planck scales.

** In English: In the strand model, a particle is a tangle of strands connected by tethers to the outer world. In Dutch: In het draadmodel is een deeltje een wirwar, die met riemen naar de buitenwereld is verbonden. En Français: Dans le modèle des fils, les particules sont des enchevêtrements connectés au monde exterieur not points, are assumed to be the fundamental constituents of vacuum, horizons, matter and radiation.

> “Nur die *ergangenen* Gedanken haben Wert.*”
> Friedrich Nietzsche

Introducing strands

The strand model starts with a simple idea:

▷ Nature is made of unobservable, fluctuating, featureless strands.

We will discover that everything observed in nature – vacuum, fermions, bosons and horizons – is made of strands. Strands are the common and extended constituents of everything. Even though strands are unobservable and featureless, all observations are due to strands.

▷ All observations, all change and all events are composed of the fundamental event, the crossing switch.

To describe all observations with precision, the strand model uses only one *fundamental principle*:

▷ ***Planck units*** are defined through **crossing switches** of strands.

Page 149 The definition of the Planck units with the crossing switch is illustrated in Figure 10. All measurements are consequence of this definition. All observations and everything that happens are composed of fundamental events. The fundamental principle thus specifies why and how Planck units are the *natural units* of nature. In particular, the four basic Planck units are associated in the following way:

▷ ***Planck's quantum of action*** $\hbar/2$ appears as the action value associated to a crossing switch. The action $\hbar$ corresponds to a double crossing switch, or full turn of one strand segment around another.*

▷ The (corrected) ***Planck length*** $l_{\mathrm{Pl}} = \sqrt{4G\hbar/c^3}$ appears as the effective diameter of strands. Since the Planck length is a limit that cannot be achieved by measurements, strands with such a diameter remain unobservable.*

▷ The ***Planck entropy***, i.e., the Boltzmann constant k, is the natural unit associated to the counting and statistics of crossings.*

▷ The (corrected) ***Planck time*** $t_{\mathrm{Pl}} = \sqrt{4G\hbar/c^5}$ appears as the shortest possible duration of a crossing switch.*

par des liens. Auf Deutsch: Im Fadenmodell sind Teilchen Gewirre aus Fäden, die durch Leinen mit der Außenwelt verbunden sind. In Italiano: Nel modello dei fili, le particelle sono degli intrecci che dei nessi collegano al mondo esterno.

* ‘Only thoughts conceived while walking have value.’ Friedrich Nietzsche (b. 1844 Röcken, d. 1900 Weimar) was philologist, philosopher and sick.

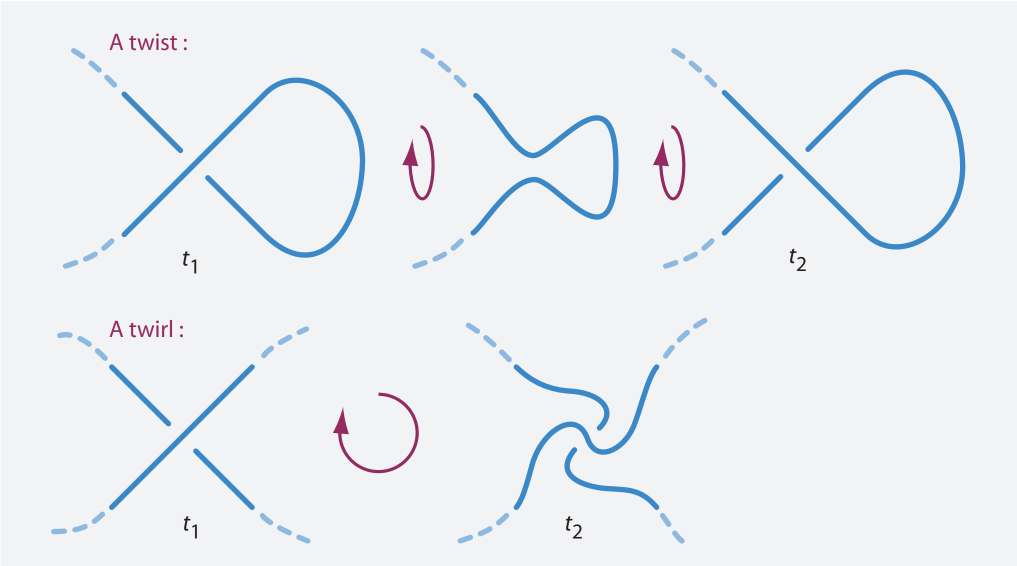


FIGURE 11 An example of strand deformation leading to a crossing switch (above) and one that does *not* lead to a crossing switch (below).

Crossing switches that are faster than the Planck time do not play a role, as they are unobservable and unmeasurable. Let us see why.

How can we imagine a minimum time interval in nature? A crossing switch could be arbitrarily fast, couldn't it? So how does the Planck time arise? To answer, we must recall the role of the observer. The observer is a physical system, also made of strands. The observer cannot define a really continuous background space-time; careful consideration tells us that the space-time defined by the observer is somewhat fuzzy: it is effectively *shivering*. The average shivering amplitude is, in the best possible case, of the order of a Planck time and length. Therefore, crossing switches faster than the Planck time are not observable by an observer made of strands.**

Strands are *impenetrable*. The switch of a crossing thus always requires the motion of strand segments *around* each other. The simplest example of a deformation leading to a crossing switch is shown in Figure 11.

Can you deduce the strand processes for the Planck momentum, the Planck force and the Planck energy? Challenge 115 s

Challenge 116 e Exploring strand processes we find: the fundamental principle implies that every Planck unit is an observer-invariant *limit value*. Therefore we get an important result: the fundamental principle naturally contains special and general relativity, quantum theory Page 37

* In other words, the strand model sets $\hbar = l_{\text{Pl}} = t_{\text{Pl}} = k = 1$. The strange numerical values that these constants have in the SI, the international system of units, follow from the strange, i.e., historical definitions of the metre, second, kilogram and kelvin.

** The issue of time remains subtle also in the strand conjecture. The requirement of consistency with macroscopic experience, realized with shivering space or space-time, allows us to side-step the issue. An alternative approach might be to picture a crossing switch and its fluctuations in 4 space-time dimensions, thus visualizing how the minimum time interval is related to minimum distance. This might be worth exploring. But also in this approach, the fuzziness due to shivering is at the basis of minimum time. Challenge 114 r

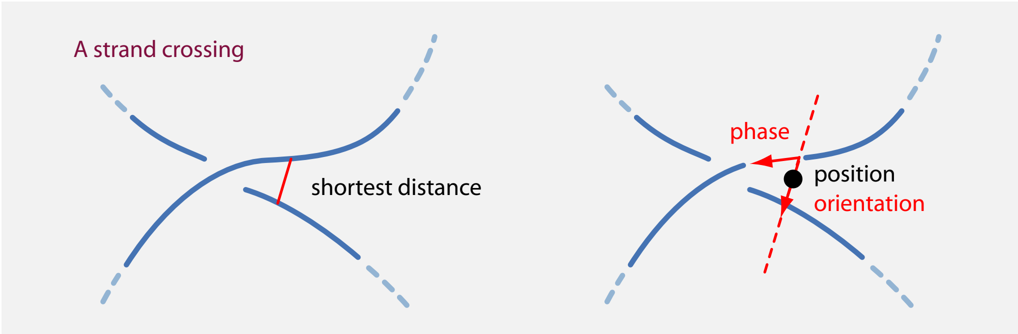


FIGURE 12 The definition of a crossing, its position, its orientation and its phase. The shortest distance defines a local density.

and thermodynamics (though seemingly not elementary particle physics, which does not arise at first sight). In theory, this argument is sufficient to show that the fundamental principle contains all these parts of twentieth century physics. In practice, however, physicists do not change their thinking habits that quickly; thus we first need to show this result in more detail. Then we continue with particle physics.

Events, processes, interactions and colours

In the strand model, every physical process is described as a sequence of crossing switches. But every physical process is also a sequence of events. We thus deduce that events are processes:

▷ Any ***event***, any observation, any measurement and any interaction is composed of switches of crossings between two strand segments.

The crossing switch is the fundamental process in nature. We will show that describing events and interactions with the help of crossing switches leads, without alternative, to the *complete* standard model of particle physics, with all its known gauge interactions, all its known particle spectrum, and all its fundamental constants.

In particular, we will show:

▷ *Particle masses,* the elementary electric charge e and the fine structure constant $\alpha = 137.036(1)$ are due to crossing switches.

The value of the fine structure constant and the standard model are *not evident* consequences of the fundamental principle. nevertheless, they are *natural* consequences – as we will find out.

From strands to modern physics

Every observation and every process is a sequence of crossing switches of unobservable strands. In turn, crossing switches are automatic consequences of the shape fluctuations of strands. We will show below that all the continuous quantities we are used to – physical space, physical time, gauge fields and wave functions – result from *averaging* crossing switches over the background space. The main conceptual tools necessary in the following are:

▷ A ***crossing*** of strands is a local minimum of strand distance. The position, orientation and phase of a crossing are defined by the space vector corresponding to the local minimum of distance, as shown in Figure 12.

Page 174 The position, orientation and phase of crossings will lead, as shown later on, to the position, orientation and phase of wave functions. The sign of the orientation is defined by arbitrarily selecting one strand as the starting strand. The even larger arbitrariness in the definition of the phase will be of great importance later on: it implies the existence of the three known gauge groups.

▷ A ***crossing switch*** is the rotation of the crossing orientation by an angle π *at a specific position*. More precisely, a crossing switch is the *inversion* of the orientation at a specific position.

We note that the definitions make use of all *three* dimensions of space; therefore the number of crossings and of crossing switches is *independent* of the direction of observation. This contrasts with the definition of crossing used in two-dimensional knot diagrams; in such two-dimensional projections, the number of crossings does depend on the direction of the projection.

We note that strand fluctuations do not conserve the number of crossings; due to fluctuations, crossings disappear and appear and disappear over time. This appearance and disappearance will turn out to be related to virtual particles.

The fundamental principle declares that events are not points on manifolds; instead,

▷ ***Events*** are (one or several) observable crossing switches of unobservable strands.

Since all observations are made of events, all experimental observations should follow Page 353 from the strand definition of an event. We will confirm this in the rest of this text. The strands are featureless: they have no mass, no tension, no stiffness, no branches, no fixed length, no ends, and they cannot be pulled, cut or pushed through each other. Strands have no measurable property at all: strands are unobservable. Only crossing switches are observable. Featureless strands are thus among the simplest possible extended constitu- Page 167 ents. How simple are they? We will discuss this issue shortly.

▷ ***Strands*** are one-dimensional curves in three-dimensional space that reach the border of space.

In practice, the *border of space* has one of two possible meanings. Whenever space is assumed to be flat, the border of space is spatial infinity. Whenever we take into account the properties of the universe as a whole, the border of space is the cosmological horizon.

Imagining the strands as having Planck diameter does not make them observable,

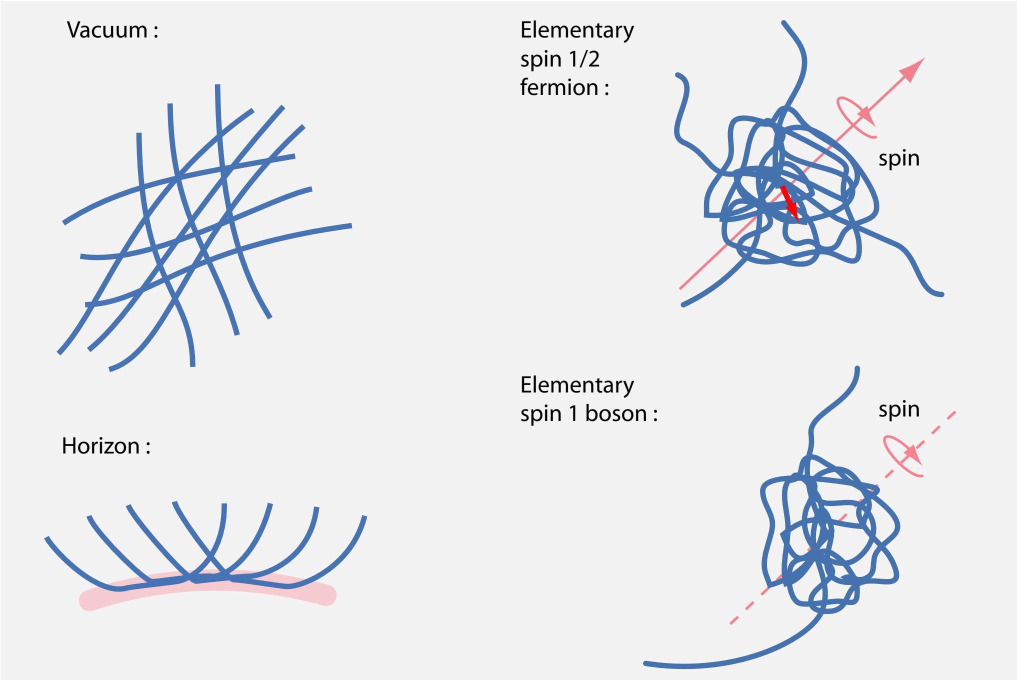


FIGURE 13 A first illustration of the basic physical systems found in nature; they will be explored in detail below.

as this measurement result cannot be realized. (We recall that the Planck length is the lower bound on any length measurement.) In low energy situations, a vanishing strand diameter is an excellent approximation.

▷ In a purist definition, featureless strands have no diameter – neither the Planck length nor zero. They are better thought as *long thin clouds*.

Strands are unobservable and featureless, and thus have no diameter. Due to shape fluctuations, or equivalently, due to the shivering of space-time, the strands can be thought as having an effective diameter, akin to the diameter of a long thin cloud; this effective diameter is just a guide to our thinking. Since it is due to the shivering of the background space-time, the strand diameter is invariant under boosts. Funnels, mentioned Page 172 below, might be a better visualization of the purist definition of strand. To keep this introduction as intuitive as possible, however, we stick with the idea of strands having an effective, invariant Planck diameter.

The strand model distinguishes *physical space* from *background space*. We will show Page 169 shortly why both concepts are required. With this distinction, the strand model asserts that matter and radiation, vacuum and horizons, are all built from *fluctuating strands* in a *continuous background*. We first clarify the two basic space concepts.

▷ ***Physical space***, or vacuum, is a physical system made of tangles that has size, curvature and other measurable properties.

TABLE 7 Correspondences between all known physical systems and mathematical tangles.

Physical system	Strand content	Tangle type
Vacuum and dark energy	many unknotted and untangled infinite strands	unlinked, trivial tangle
Graviton	two infinite twisted strands, asymptotically parallel	rational tangle
Gravity wave	many infinite twisted strands	many rational tangles
Horizon	many woven infinite strands	a weave: a woven, web-like tangle
Photon (radiation)	one unknotted helical strand	trivial 1-tangle
Classical electromagnetic wave (radiation)	many unknotted and untangled helical strands	many helically deformed/tangled curves
Weak vector boson before symmetry breaking (radiation)	two locally curved, asymptotically parallel strands	trivial rational 2-tangle
W and Z boson	three infinite, slightly linked strands, asymptotically in a plane	rational 3-tangle
Gluon (radiation)	three infinite locally curved, asymptotically parallel strands	trivial rational 3-tangle
Elementary quark (matter)	*two* infinite linked strands	rational 2-tangle
Elementary lepton (matter)	*three* infinite linked strands	rational 3-tangle
Higgs boson	*three* infinite linked strands	braided 3-tangle

▷ ***Continuous background space*** is introduced by the observer only to be able to describe observations. Every observer introduces his own background. It does not need to coincide with physical space, and it does not do so at the location of matter or black holes. But every observer's background is continuous and has three spatial and one temporal dimension.

Page 209

At this point of the discussion, we simply *assume* background space. Later on we will see why background space appears and why it *needs* to be three-dimensional. The *size* of the background space is assumed to be large; larger than any physical scale under discussion. In most situations of everyday life, when space is flat, background space and physical space coincide. However, they differ in situations with curvature and at Planck energy.

▷ ***Fluctuations*** change the position, shape and length of strands; fluctuations thus change position, orientation and phase of strand crossings. However, fluctuations never allow one strand to pass through another.

All strand fluctuations are possible, as long as strands do not interpenetrate. For example, there is no speed limit for strands. Whenever strand fluctuations lead to a crossing switch, they lead to an observable effect – be it a vacuum fluctuation, a particle reaction or a horizon fluctuation.

▷ Fluctuations are a consequence of the embedding of strands in a continuous background.

In the strand model, even isolated physical systems are surrounded by a bath of fluctuating vacuum strands. The properties of fluctuations, such as their spectrum, their density etc., are fixed once and for all by the embedding. Fluctuations are necessary for the self-consistency of the strand model.

Due to the impenetrability of strands – which itself is a consequence of the embedding in a continuous background – any disturbance of the vacuum strands at one location *propagates*. We will see below what disturbances exist and how they differ from fluctuations.

Fluctuating strands that lead to crossing switches explain everything that *does* happen, and explain everything that does *not* happen. Our main aim in the following is to classify all possible strand fluctuations and all possible strand configurations, in particular, all states that differ from flat vacuum states. By doing so, we will be able to classify every process and every system that we observe in nature.

We will discover that *all physical systems* can be constructed from strands. Table 7 gives a first overview of how vacuum, particles and horizons result from *tangles* of strands.

▷ A ***tangle*** is a configuration of one or more strands that are *linked* or *knotted*. Tangles are characterized by their topology, i.e., by the precise way that they are linked or knotted.

Page 156 Some examples of important tangles are given in Figure 13. They will be discussed in detail in the following. Among others, we will discover that knots and knotted tangles do *not* play a role in the strand conjecture; only linked, but unknotted tangles do.

We observe that vacuum, matter and radiation are all made of the *same* fundamental Page 313 constituents, as required for a complete theory. We will discover below that classifying localized tangles naturally leads to the elementary particles that make up the standard model of particle physics – and to no other elementary particle.

We will also discover that strand fluctuations and the induced crossing switches in every physical system lead to the evolution equations and the Lagrangians of quantum field theory and of general relativity. In this way, strands describe *every* physical process observed in nature, including *the four known interactions*, and every type of motion.

Page 152 The fundamental principle relates crossing switches and observations. The fundamental principle was discovered because it appears to be the only simple definition of Planck units that on the one hand yields space-time, with its continuity, local isotropy and curvature, and on the other hand realizes the known connection between the Ref. 154 quantum of action, spin and rotation.

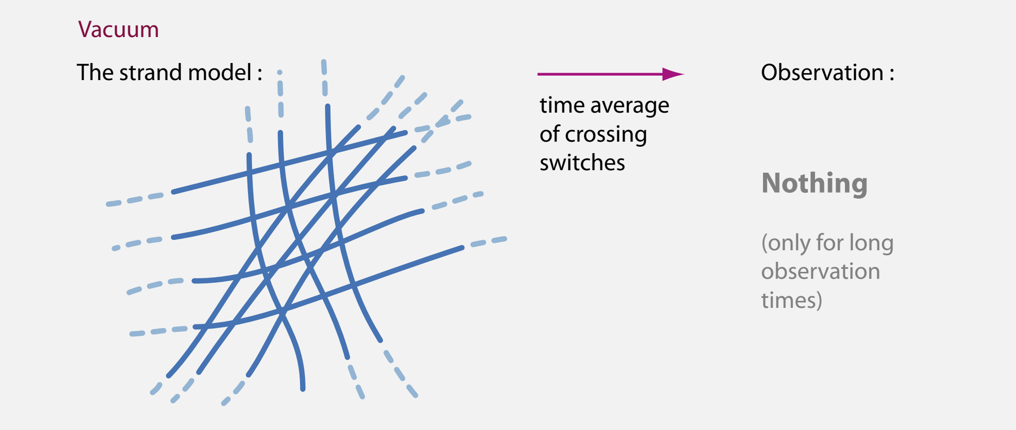


FIGURE 14 An illustration of the strand model for the vacuum.

Vacuum

We now construct, step by step, all important physical systems, concepts and processes from tangles. We start with the most important.

▷ ***Vacuum***, or ***physical space***, is formed by the time average of many unknotted fluctuating strands.

Figure 14 visualizes the definition. In the following, vacuum and physical space are always taken to be synonyms; the exploration will show that this is the most sensible use of the two concepts.* However, as mentioned, the strand model distinguishes *physical* space from *background* space. In particular, since matter and vacuum are made of the same constituents, it is impossible to speak of physical space at the location of matter. At the location of matter, it is only possible to use the concept of background space.

When the strand fluctuations in flat vacuum are averaged over time, there are no crossing switches. Equivalently, if we use concepts to be introduced shortly, flat vacuum shows, *averaged over time*, no tangles and no crossing switches, so that it is observed to be empty of matter and radiation. *Temporary* tangles that appear for a short time through vacuum fluctuations will be shown below to represent virtual particles. Page 239

We note that the (flat) physical vacuum state, which appears after averaging the strand crossings, is *continuous*, *Lorentz invariant* and *unique*. These are important points for the consistency of the model. Later we will also discover that curvature and horizons have a natural description in terms of strands; exploring them will yield the field equations of general relativity. Page 281

We also note that Figure 14 implies, despite appearances, that vacuum is *isotropic*. To see this, we need to recall that the observables are the crossing switches, not the strands, and that the observed vacuum isotropy results from the isotropy of the time-averaged strand fluctuations.

* We recall that since over a century, the concept of aether is superfluous, because it is indistinguishable from the concept of vacuum.

▷ We do not make any statement on the numerical density of strands in vacuum, or, equivalently, on their average spacing. Since strands are not observable, such a statement is not sensible. In particular, strands in vacuum are *never* tightly packed.

With the definition of the vacuum as a time average, the strand conjecture yields a *minimum length* and a *continuous vacuum* at the same time. In this way, many issues about the alleged contradiction between continuity and minimum length are put to rest. In particular, physical space is *not* fundamentally discrete: a minimum length appears, though only in domains where physical space is undefined. On the other hand, the continuity of physical space results from an averaging process. Therefore, physical space is *not* fundamentally continuous: the strand model describes physical space as a homogeneous distribution of crossing switches. This is the strand version of Wheeler's idea space-time foam.

The strand model thus replaces what used to be called 'space-time foam' or 'quantum foam' with a web of strands. The biggest difference between the two ideas is the lack of any topology change in the strand model. Space, and in particular background space, is always three-dimensional.

Observable values and limits

The fundamental principle implies the following definitions of the basic observables:

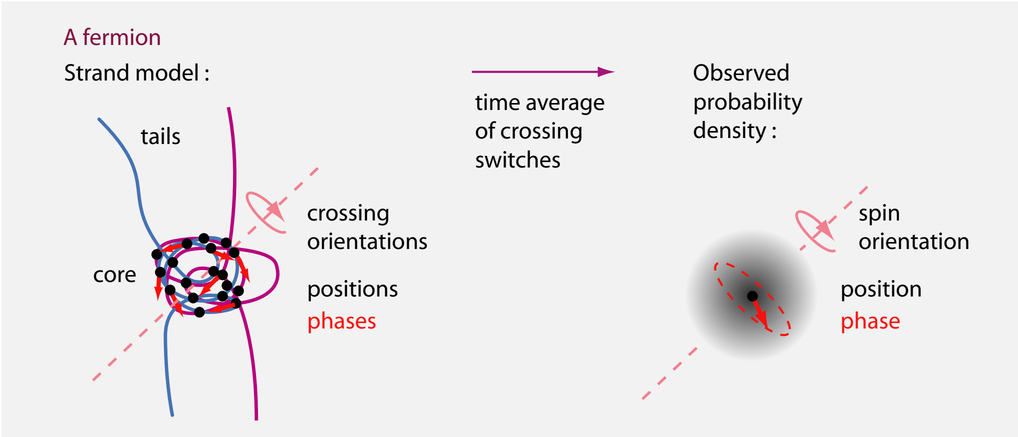
▷ The ***distance*** between two particles is the maximum number of crossing switches that can be measured between them. Length measurement is thus defined as counting Planck lengths.

▷ The ***time interval*** between two events is the maximum number of crossing switches that can be measured between them. Time measurement is thus defined as counting Planck times.

▷ The physical ***action*** of a physical system evolving from an initial to a final state is the number of crossing switches that can be measured. Action measurement is thus defined as counting crossing switches. Physical action is thus a measure for the *change* that a system undergoes.

▷ The ***entropy*** of any physical system is related to the logarithm of the number of possible measurable crossing switches. Entropy measurement is thus defined through the counting of crossing switches. The strand model thus states that *any large physical system* – be it made of matter, radiation, empty space or horizons – has entropy.

It is well-known that all other physical observables are defined using these four basic ones. In other words, *all* physical observables are defined with crossing switches. We also note that even though counting always yields an integer, the result of a physical measurement is often an *average* of many counting processes. As a result of averaging and fluctuations, measured values can be non-integer multiples of Planck units. Therefore, space, time, action, entropy and all other observables are *effectively* real numbers, and



Page 174 FIGURE 15 The tangle model of a spin 1/2 particle. More details will be given below.

thus continuous. Continuity is thus reconciled through averaging with the existence of a minimum measurable length and time interval.

Finally, we note that defining observables with the help of crossing switches automatically makes the Planck units c, $\hbar$, $c^4/4G$, k and all their combinations both *observer-invariant* and *limit* values. All these conclusions agree with the corresponding requirements for a final theory of nature. Crossing switches thus explain the origin of these limits. This solves two issues in the millennium list of Table 1.19 An pretentious summary would be: strands explain the (new) SI system of units.

Page 149

Particles and fields

Strands also define particles, as illustrated in Figure 15:

▷ A ***quantum particle*** is a *tangle* of fluctuating strands. The tangle *core*, the region where the strands are linked, defines position, speed, phase and spin of the particle. The tangle *tails* reach up to the border of space.

Page 176 As shown in more detail soon, this definition of quantum particles yields, depending on the tangle details, either fermion or boson behaviour, and reproduces the spin–statistics theorem.

Boson tangles will allow us to model field intensities. In particular, boson tangles allow us to deduce the electromagnetic and the two nuclear fields, as well as the corresponding gauge symmetries of the standard model of particle physics. Page 225

Page 215 Modelling fermions as tangles will allow us to deduce Dirac's equation for relativistic quantum particles (and the Schrödinger equation for non-relativistic particles). Still later, by classifying all possible tangles, we will discover that only a *finite* number of possible elementary particles exist, and that the *topological type* of tangle determines the mass, mixings, quantum numbers, charges and couplings of each elementary particle. We can also speak of the *tangle model* of particles. Page 314

In the 1960s, John Wheeler stated that a unified description of nature must explain 'mass without mass, charge without charge, field without field'. The strand conjecture realizes this aim, as we will find out.

Before we deduce modern physics, we first take a break and explore some general issues of the strand conjecture.

Curiosities and fun challenges about strands

Why do crossing switches have such a central role in the strand model? An intuitive explanation follows from their role in the definition of observables. All measurements – be they measurements of position, speed, mass or any other observable – are electromagnetic. In other words, all measurements in nature are, in the end, detection of photons. And the strand model shows that photon absorption and detection are intimately related to the crossing switch, as we will find out below. Page 228

* *

Is there a limit to the fluctuations of strands? Yes and no. On the one hand, the 'speed' of fluctuations is unlimited. On the other hand, fluctuations with a 'curvature radius' smaller than a Planck length do not lead to observable effects. Note that the terms 'speed' and 'radius' are between quotation marks because they are unobservable. Care is needed when talking about strands and their fluctuations.

* *

What are strands made of? This question tests whether we are really able to maintain the fundamental circularity of the unified description. Strands are featureless. They have no measurable properties: they have no branches, carry no fields and, in particular, they cannot be divided into parts. The 'substance' that strands are made of has no properties. Thus strands are not made of anything. This may seem surprising at first. Strands are extended, and we naturally imagine them as sequence of points. But this is a fallacy. Given the way that observations and events are defined, there is no way to observe, to label or to distinguish points on strands. Crossing switches do not allow doing so, as is easily checked: the mathematical points we imagine on a strand are not physical points. 'Points' on strands are unobservable: they simply do not exist. Challenge 117 e

But strands must be made of something, we might insist. Later we will find out that in the strand model, the universe is made of a single strand tangled in a complicated way. *Nature is one strand.* Therefore, strands are not made of something, they are made of everything. The substance of strands is nature itself. Page 306

* *

Since there is only one strand in nature, strands are not a reductionist approach. At Planck scale, nature is one. And indivisible.

* *

What are elementary particles? In the strand model, elementary particles are not elementary. Elementary particles are (families of) *tangles* of strands. In other words, elementary particles are not the basic building blocks of matter – strands are. If particles could *really* be elementary, it would be impossible to understand their properties.

In the strand conjecture, elementary particles are not really elementary, but neither are they, in the usual sense, composed. Particles are tangles of unobservable strands. In this way, the strand conjecture retains the useful aspects of the idea of elementary particle but gets rid of its limitations. In a sense, the strand model can be seen as eliminating the concepts of elementariness and of particle. In the same way, strands get rid of the idea of point in space. This confirms and realizes another requirement for a complete description that we had deduced earlier on. Page 78

* *

Challenge 118 e Can macroscopic determinism arise at all from randomly fluctuating strands?

* *

Challenge 119 s Do parallel strands form a crossing? Do two distant strands form a crossing?

* *

Challenge 120 s Is a crossing switch defined in more than three dimensions?

* *

Challenge 121 s Can you find a way to generalize or to modify the strand conjecture?

* *

In hindsight, the fundamental principle resembles John Wheeler's vision 'it from bit'. He formulated it, among others, in 1989 in his often-cited essay *Information, physics, quantum: the search for links.*

* *

Looking back, we might equally note a relation between the strand conjecture and the expression 'it from qubit' that is propagated by David Deutsch. A qubit is a quantum-mechanical two-level system. What is the difference between the fundamental principle and a qubit? Challenge 122 e

* *

Is the strand model confirmed by other, independent research? Yes, a few years after the strand model appeared, this started to happen. For example, in a long article exploring the small scale structure of space-time from various different research perspectives in general relativity, Steven Carlip comes to the conclusion that all these perspectives suggest the common result that 'space at a fixed time is thus threaded by rapidly fluctuating lines'. This is exactly what the strand model states. Ref. 155

Other theoretical approaches that confirm the strand model are mentioned in various places later in the text. Page 301 Despite such developments, the essential point remains to check how the strand model compares with experiment. Given that the strand model turns out to be unmodifiable, there are no ways to amend predictions that turn out to be wrong. If a single prediction of the strand model turns out to be incorrect, the model is doomed. But so far, no experimental prediction of the strand model contradicts experiments. Page 412

Do strands unify? – The millennium list of open issues

Does the strand conjecture reproduce all the paradoxical results we found in the first chapters? Yes, it does. The strand conjecture implies that vacuum cannot be distinguished from matter at Planck scales: both are made of strands. The strand conjecture implies that observables are not real numbers at Planck scales. The strand conjecture implies that the universe and the vacuum are the same, when explored at high precision: both are made of one strand. The strand conjecture also implies that the number of particles in the universe is not clearly defined and that nature is not a set. You can check by yourself that all other paradoxes appear automatically. Furthermore, almost all requirements for a final theory listed in Table 6 are fulfilled. Only two requirements of the table must be discussed in more detail: the requirements of complete precision and of unmodifiability. We start with complete precision.

Page 65 Page 74 Page 91 Page 129 Challenge 123 e

If strands really describe *all* of nature, they must explain the inverse square dependence with distance of the electrostatic and of the gravitational interaction. But that is not sufficient. If the strand conjecture is a complete, unified description, it must provide *complete* precision. This requires, first of all, that the conjecture describes *all* experiments. As will be shown below, this is indeed the case, because the strand conjecture contains both general relativity and the standard model of particle physics. But secondly and most importantly, the conjecture must also settle all those questions that were left unanswered by twentieth-century fundamental physics. Because the questions, the *millennium list* of open issues, are so important, they are given, again, in Table 8.

TABLE 8 The millennium list: *everything* the standard model and general relativity *cannot* explain; thus, also the list of the *only* experimental data available to test the final, unified description of motion.

Observable	Property unexplained since the year 2000
Local quantities unexplained by the standard model: particle properties	
$\alpha = 1/137.036(1)$	the low energy value of the electromagnetic coupling or fine structure constant
α_w or θ_w	the low energy value of the weak coupling constant or the value of the weak mixing angle
α_s	the value of the strong coupling constant at one specific energy value
m_q	the values of the 6 quark masses
m_l	the values of 6 lepton masses
m_W	the value of the mass of the W vector boson
m_H	the value of the mass of the scalar Higgs boson
$\theta_{12}, \theta_{13}, \theta_{23}$	the value of the three quark mixing angles
δ	the value of the CP violating phase for quarks
$\theta^\nu_{12}, \theta^\nu_{13}, \theta^\nu_{23}$	the value of the three neutrino mixing angles
$\delta^\nu, \alpha_1, \alpha_2$	the value of the three CP violating phases for neutrinos
$3 \cdot 4$	the number of fermion generations and of particles in each generation
J, P, C, etc.	the origin of all quantum numbers of each fermion and each boson
Concepts unexplained by the standard model	
$c, \hbar, k$	the origin of the invariant Planck units of quantum field theory

TABLE 8 (Continued) The millennium list: *everything* the standard model and general relativity *cannot* explain; also the *only* experimental data available to test the final, unified description of motion.

Observable	Property unexplained since the year 2000
$3 + 1$	the number of dimensions of physical space and time
SO(3,1)	the origin of Poincaré symmetry, i.e., of spin, position, energy, momentum
Ψ	the origin and nature of wave functions
$S(n)$	the origin of particle identity, i.e., of permutation symmetry
Gauge symmetry	the origin of the gauge groups, in particular:
U(1)	the origin of the electromagnetic gauge group, i.e., of the quantization of electric charge, of the vanishing of magnetic charge, and of minimal coupling
SU(2)	the origin of weak interaction gauge group, its breaking and P violation
SU(3)	the origin of strong interaction gauge group and its CP conservation
Renorm. group	the origin of renormalization properties
$\delta W = 0$	the origin of the least action principle in quantum theory
$W = \int L_{\mathrm{SM}}\, \mathrm{d}t$	the origin of the Lagrangian of the standard model of particle physics
Global quantities unexplained by general relativity and cosmology	
0	the observed flatness, i.e., vanishing curvature, of the universe
$1.2(1) \cdot 10^{26}$ m	the distance of the horizon, i.e., the 'size' of the universe (if it makes sense)
$\rho_{\mathrm{de}} = \Lambda c^4/(8\pi G) \approx 0.5\,\mathrm{nJ/m^3}$	the value and nature of the observed vacuum energy density, dark energy or cosmological constant
$(5 \pm 4) \cdot 10^{79}$	the number of baryons in the universe (if it makes sense), i.e., the average visible matter density in the universe
ρ_{dm}	the density and nature of dark matter
$f_0(1, ..., c.\ 10^{90})$	the initial conditions for *c.* 10^{90} particle fields in the universe (if or as long as they make sense), including the homogeneity and isotropy of matter distribution, and the density fluctuations at the origin of galaxies
Concepts unexplained by general relativity and cosmology	
c, G	the origin of the invariant Planck units of general relativity
$\mathrm{R} \times \mathrm{S}^3$	the observed topology of the universe
$G^{\mu\nu}$	the origin and nature of curvature, the metric and horizons
$\delta W = 0$	the origin of the least action principle in general relativity
$W = \int L_{\mathrm{GR}} \mathrm{d}t$	the origin of the Lagrangian of general relativity

The open issues in the millennium list must be resolved by any complete, unified description of nature, and thus also by the strand model. All the open issues of the list can be summarized in two general requests:

— *Reproduce* quantum theory, the standard model, general relativity and cosmology.

— *Explain* masses, mixing angles and coupling constants.

Of course, only the second point is the *definite test* for a complete, unified description. But we need the first point as well. The following chapters explore both points.

Are strands final? – On generalizations and modifications

> The chief attraction of the theory lies in its logical completeness. If a single one of the conclusions drawn from it proves wrong, it must be given up; to modify it without destroying the whole structure seems to be impossible.
>
> Albert Einstein, *The Times*, 28. 11. 1919.

If a description of motion claims to be *complete*, it must explain *all* aspects of motion. To be a full explanation, such a description must not only be logically and experimentally complete, it must also be *unmodifiable*. Even though Einstein made the point for general relativity, this important aspect merits a few additional remarks. In particular, any unmodifiable explanation has two main properties: first, it cannot be generalized, and second, it is not itself a generalization.

Generalizing models is a sport among theoretical and mathematical physicists. If you have a description of a part of nature, they will try to find more general cases. For any candidate unified description, they will try to explore the model in more than three dimensions, with more than three generations of quarks, with more complicated gauge symmetries, with different types of supersymmetry, with more Higgs bosons, or with additional heavy neutrinos. In the case of the strand model, researchers will also explore models with more complicated entities than strands, such as bands or bifurcating entities, and any other generalization they can imagine.

▷ Can a *final* description of nature have generalizations? No.

Indeed, if it were possible to generalize the complete description, it would lose the ability to *explain* any of the millennium issues! If a candidate unified theory could be generalized, it would not be final or complete. In short, if the strand model is a complete description, the efforts of theoretical and mathematical physicists just described must all be impossible. So far, investigations confirm this prediction: no generalization of the strand model has been found yet.

Where does this fondness for generalization come from? In the history of physics, generalizations often led to advances and discoveries. In the past, generalizations often led to descriptions that had a *wider range* of validity. As a result, generalizing became the way to search for new discoveries. Indeed, in the history of physics, the old theory often was a *special case* of the new theory. This relation was so common that usually, *approximation* and *special case* were taken to be synonyms. This connection leads to a second point.

General relativity and the standard model of particle physics must indeed be *approximations* of the final theory. But can either general relativity or the standard model be *special cases* of the complete, unified theory? Or, equivalently:

▷ Can the unified theory be a generalization of existing theories? No.

Because neither general relativity nor the standard model of particle physics are able to explain the millennium issues, any generalization of them would also be unable to do so. Generalizations have *no* explanatory power. If the unified theory were a generalization of the two existing theories, it could not explain any of the millennium issues of Table 8! Therefore, general relativity and the standard model of particle physics must be approximations, but *not* special cases, of the final theory.* In particular, if the strand model is a complete description, *approximations* of the strand model must exist, but *special cases* must not. This is indeed the case, as we will find out.

To summarize, a unified theory must be an *explanation* of all observations. An explanation of an observation is the recognition that it follows unambiguously, *without alternative*, from a general property of nature. We conclude that the complete, unified description of motion must neither allow generalization nor must it be a generalization of either the standard model or general relativity. The unified theory cannot be generalized and cannot be 'specialized'; the unified theory must be *unmodifiable*.** This requirement is extremely strong; you may check that it eliminates most past attempts at Challenge 124 e unification. For example, this requirement eliminates grand unification, supersymmetry and higher dimensions as aspects of the final theory: indeed, these ideas *generalize* the standard model of elementary particles and they are *modifiable*. Therefore, all these ideas lack explanatory power.

A complete and unified theory must be an *unmodifiable explanation* of general relativity and the standard model. Because neither supersymmetry, nor the superstring conjecture, nor loop quantum gravity explain the standard model of particle physics, they are not unified theories. Because these models are modifiable, they are not complete. In fact, at least one of these two aspects is lacking in every candidate unified theory proposed in the twentieth century.

We will discover below that *the strand conjecture is unmodifiable*. Its fundamental Challenge 125 e principle cannot be varied in any way without destroying the whole description. Indeed, no modification of the strand conjecture or of the fundamental principle has been found so far. We will also discover that the strand conjecture *explains* the standard model of particle physics and explains general relativity. The strand conjecture is thus a candidate for the final theory.

Why strands? – Simplicity

> “Simplex sigillum veri.***
> Antiquity ”

Let us assume that we do not know yet whether the strand conjecture can be modified or not. Two other reasons still induce us to explore featureless strands as basis for a unified description. First, featureless strands are the *simplest* known model that unifies quantum field theory and general relativity. Second, featureless strands are the *only*

* Indeed, we already found out above151 that a complete theory must contradict both quantum theory and general relativity, in particular because of the use of extended constituents.

Ref. 156 ** Independently, David Deutsch made a similar point with his criterion that an explanation is only correct if it is *hard to vary*. Used in the case of a final theory, we can say that the final theory must be an *explanation* of general relativity and of the standard model. This implies that the final theory must be hard to vary. This matches the above conclusion that the final theory must be unmodifiable.

*** 'Simplicity is the seal of truth.'

Page 106 *known* model that realizes an important requirement: a unified description must not be based on points, sets or any axiomatic system. Let us explore the issue of simplicity first.

In order to reproduce three-dimensional space, Planck units, spin, and black-hole entropy, the fundamental constituents must be *extended* and *fluctuating*. We have deduced Page 117 this result in detail in the previous chapter. The extension must be one-dimensional, because this is the simplest option, and it is also the only option compatible with three-dimensional space. In fact, one-dimensional strands explain the three-dimensionality of space, because tangles of one-dimensional strands exist *only* in three spatial dimensions. In four or more dimensions, any tangle or knot can be undone; this is impossible in three spatial dimensions.

Ref. 140 Ref. 157 Ref. 158 Ref. 159 No *simpler* model than featureless strands is possible. All other extended constituents that have been explored – ribbons, bands, strings, membranes, posets, branched lines, networks, crystals and quantum knots – increase the complexity of the model. In fact these constituents increase the complexity in *two* ways: they increase the number of features of the fundamental constituents and they complicate the mapping from the model to observation.

First, no other model based on extension uses *featureless* constituents. In all other models, the fundamental constituents have properties such as tension, field values, coordinates, quantum numbers, shape, twists, orientation, non-trivial topological information, etc. In some models, space-time is non-commutative or fermionic. All these features are *assumed*; they are added to the model by fiat. As such, they allow alternatives and are difficult if not impossible to justify. In addition, these features increase the complexity of the possible processes. In contrast, the strand model has no justification issue and no complexity issue.

Secondly, the link between more complicated models and experiment is often *intricate* and sometimes not unique. As an example, the difficulties to relate superstrings to Ref. 150 experiments are well-known. In contrast, the strand conjecture argues that the experimentally accessible Dirac equation of quantum field theory and the experimentally accessible field equations of general relativity arise *directly*, from an averaging procedure of crossing switches. Indeed, the strand conjecture proposes to unify these two halves of physics with only one fundamental principle: strand crossing switches define Planck units. In fact, we will find out that the strand conjecture describes not only vacuum and matter, but also gauge interactions and particle properties as *natural* consequences of the structure of nature at Planck scales. The comparable ideas in other models are much more elaborate.

We remark that building three-dimensional physical space from strands is even simpler than building it from points! In order to build three-dimensional space from *points*, we need concepts such as sets, neighbourhoods, topological structures and metric structures. And despite all these intricate concepts, the concept of space defined in this way still has no defined physical length scale; in short, it is *not* the same as physical space. In contrast, in order to build three-dimensional physical space from *strands*, we need no fundamental points, sets, or metric structures; we only need long-time averages of strands and their crossings. And the length scale is built in.

All this suggests that the strand model, based on featureless, one-dimensional and fluctuating constituents, might be the model for unification with the smallest number of concepts, thus satisfying Occam's razor. In fact, we will discover that strands indeed

TABLE 9 The differences between nature and any description.

Nature	Description
Nature is not a set.	Descriptions need sets to allow talking and thinking.
Nature has no events, no points and no continuity.	Descriptions need events, points and continuous 3 + 1-dimensional space-time to allow formulating them.
Nature has no point particles.	Descriptions need point particles to allow talking and thinking.
Nature is not local.	Descriptions need locality to allow talking and thinking.
Nature has no background.	Descriptions need a background to allow talking and thinking.
Nature shows something akin to $R \leftrightarrow 1/R$ duality.	Descriptions need to break duality to allow talking and thinking.
Nature is not axiomatic but contains circular definitions.	Axiomatic descriptions are needed for precise talking and thinking.

are the *simplest* way to model particles, interactions and the vacuum, while fulfilling the requirements of a final theory.

The simplicity of a model helps in two ways. First, the simpler a model is, the freer it is of ideology, preconceptions and beliefs. Secondly, the simpler a model is, the easier it can be checked against observation. In particular, a simple model allows simple checking of its solution of paradoxes. Above all, we can resolve the most important paradox of physics.

Why strands? – The fundamental circularity of physics

> “Without the concepts *place*, *void* and *time*, change cannot be. [...] It is therefore clear [...] that their investigation has to be carried out, by studying each of them separately.”
>
> Aristotle *Physics*, Book III, part 1.

The strand model describes strands as fluctuating in a background space-time of three plus one space-time dimensions. The background space-time is introduced by the observer. The background is thus different for every observer; however, all such backgrounds have three dimensions of space and one of time. The observer – be it a machine, an animal or a human – is itself made of strands, so that in fact, the background space is itself the product of strands.

We therefore have a fundamental circular definition: we describe strands with a background, and the background with strands. *Strands thus do not provide an axiomatic system in the mathematical sense.* This fulfils one of the requirements for the unified
Page 110
description.

Why does the fundamental circular definition arise? Physics is talking (and thinking) about nature and motion. A unified model of physics is talking about motion with highest precision. This implies that on the one hand, as talkers, we must use *concepts that allow us to talk*. Talking and thinking requires that we use continuous space and time: on short, we must use a *background*. The background must be continuous, without minimum length. On the other hand, to talk *with precision*, we must have a minimum length, and use *strands*. There is no way to get rid of this double and apparently contradictory requirement. More such contradictory requirements are listed in Table 9. We know that nature is not a set, has no points, no point particles and no locality, but that it is dual. But in order to talk about nature, we need a background that lacks all these properties. Because there is no way to get rid of these apparently contradictory requirements, we don't:

▷ We use both continuous background space-time and discrete strands to describe nature.

In a few words: A unified model of physics allows talking about motion with highest precision; this requirement forces us to use, *at the same time*, both continuous space-time and discrete strands. This double use is not a contradiction but, as just explained, the result of a *circular definition*. Since we, the talkers, are part of nature, a unified model means that we, the talkers, talk about ourselves.

We stress that despite the circularity of physics, Gödel's incompleteness theorem does not apply to the situation. In fact, the theorem does not apply to any unified theory of physics for two reasons. First, the incompleteness theorem applies to *self-referential* statements, not to circular definitions. Self-referential statements do not appear in physics, not in sensible mathematics and not in the strand model. Secondly, Gödel's theorem applies to mathematical structures based on sets, and the unified theory is not based on sets.

Vol. III, page 309

We do not state that *background* space and time exist *a priori*, as Immanuel Kant states, but only that background space and time are *necessary* for thinking and talking, as Aristotle states. In fact, *physical* space and time result from strands, and thus do not exist a priori; however, background space and time are required concepts for any description of observations, and thus necessary for thinking and talking. Figure 16 illustrates the solution proposed by the strand model.

Ref. 160

We have always to be careful to keep the fundamental circular definition of strands and backgrounds *in our mind*. Any temptation to resolve it leads astray. For example, if we attempt to define *sets* or *elements* (or points) with the help of measurements, we are hiding or forgetting the fundamental circularity. Indeed, many physicists constructed and still construct axiomatic systems for their field. The fundamental circularity implies that axiomatic systems are possible for *parts* of physics, but not for physics as a whole. Indeed, there are axiomatic descriptions of classical mechanics, of electrodynamics, of quantum theory, of quantum field theory, and even of general relativity. But there is no axiomatic system for *all* of physics – i.e., for the description of all motion – and there cannot be one.

Page 110

A further issue must be discussed in this context. As mentioned, strands fluctuate in a background space, and only crossing switches can be observed. In particular, this

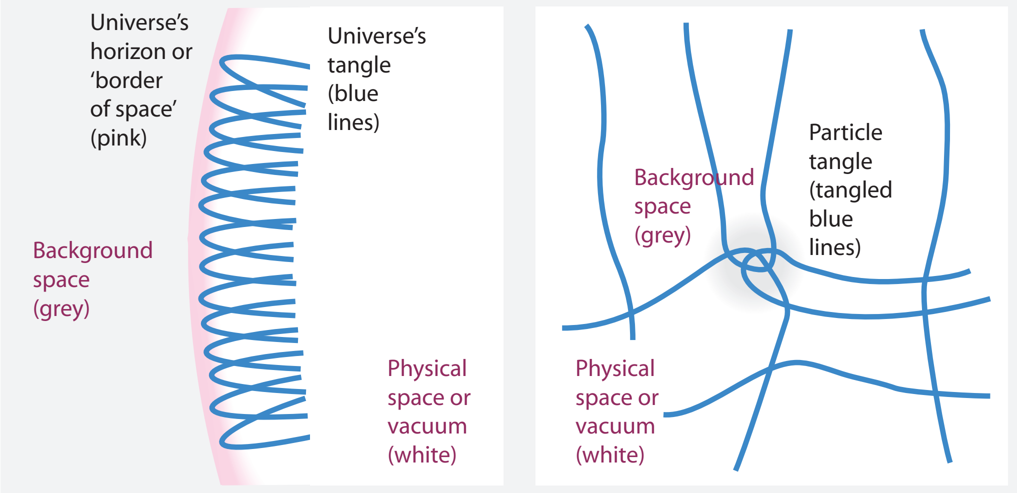


FIGURE 16 In the strand conjecture, physical space – or vacuum – and background space are distinct, both near the horizon and near particles.

implies that the mathematical points of the background space cannot be observed. In other words, despite using mathematical points to describe the background space (and strands themselves), none of them have physical significance. *Physical points do not exist in the strand model.* Physical locations of events are due to crossing switches, and can at best be localized to within a Planck length. The same limitation applies to physical events and to physical locations in time. *A natural Planck-scale non-locality is built into the model.* This realizes a further requirement that any unified description has to fulfil.

The situation for physicists working on unification is thus harder – and more fascinating – than that for biologists, for example. Biology is talking about living systems. Biologists are themselves living systems. But in the case of biologists, this does *not* lead to a circular definition. Biology does not use concepts that contain circular definitions: a living being has no problems describing *other* living beings. Even neurobiologists, who aim to explore the functioning of the brain, face no fundamental limit doing so, even though they explore the human brain using their own brain: a brain has no problem describing other brains. In contrast, physicists working on unification need to live with circularity: a fundamental, precise description of motion requires to be conscious of our own limitations as describing beings. And our main limitation is that we cannot think without continuous space and time, even though these concepts do not apply to nature.

Ref. 161

We conclude: *A unified description cannot be axiomatic, cannot be based on observable physical points, must distinguish physical space from background space, and cannot be background-independent.* Many models based on extended constituents also use backgrounds. However, most models also allow the definition of sets and axiomatic descriptions. Such models thus cannot be candidates for a unified description of nature. In contrast, the strand conjecture keeps the fundamental circularity of physics intact; it does not allow an axiomatic formulation of fundamental physics, and only allows points or sets as approximate concepts.

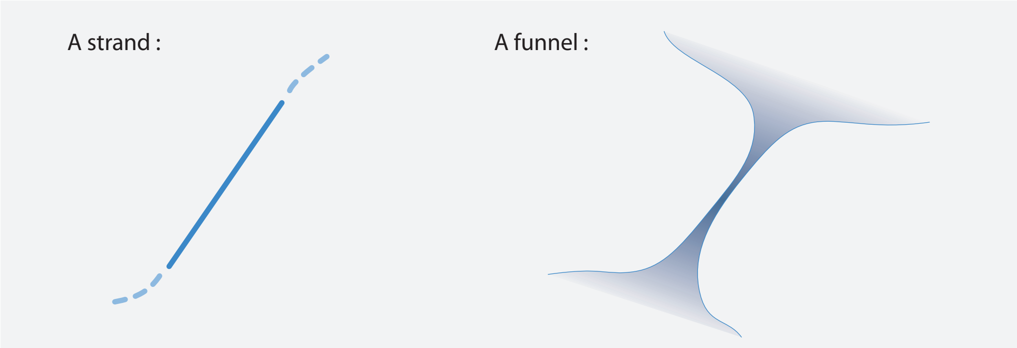


FIGURE 17 Two equivalent depictions of the fundamental constituents of nature: strands and funnels.

Funnels – an equivalent alternative to strands

Another type of constituent also fulfils all the conditions for a unified description. As shown in Figure 17, as an alternative to fluctuating strands, we can use fluctuating *funnels* as fundamental constituents. In the description with funnels, nature resembles a complicated tangle of a three-dimensional space that is projected back into three dimensions.

Funnels show that the strand model only requires that the effective *minimal* effective diameter of a strand is the Planck length; it could have other diameters as well. Funnels also show that due to varying diameters, strands can, through their fluctuations, literally be everywhere in space and thus effectively *fill* space, even if their actual density is low.

Funnels resemble many other research topics. Funnels are similar to wormholes; however, both their ends lead, at the border of space, 'into' usual three-dimensional space. Funnels are also similar to D-branes, except that they are embedded in three spatial dimensions, not ten. Funnels also resemble a part of an exotic manifold projected into three-dimensional space. Fluctuating funnels also remind us of the amoeba mentioned Page 131 above. However, the similarities with wormholes, D-branes or exotic manifolds are of little help: so far, none of these approaches has led to viable models of unification.

A first check shows that the funnel alternative seems *completely equivalent* to strands.* You might enjoy testing that all the conclusions deduced in the following pages appear Challenge 126 e unchanged if strands are replaced by funnels. In particular, also funnels allow us to deduce quantum field theory, the standard model and general relativity. Due to the strict equivalence between strands and funnels, the choice between the two alternatives is a matter of taste or of visualization, but not a matter of physics. We use strands in the following, as they are simpler to draw.

Knots and the ends of strands

In the original strand model, developed in the year 2008, strands that contain knots were part of the allowed configurations. This has the disadvantage that the formation of a knot requires at least one loose end that is pulled through a strand configuration. Such loose

* Two issues that put this equivalence into question are ending funnels and diameter behaviour under boosts. The first issue is subject of research, but it is expected that it poses no problem. The second issue is mitigated by the shivering of the background space.

ends, however, produce a number of issues that are difficult to explain, especially during the emission and absorption of knotted tangles.

Later, in 2015, it became clear that strands *without* knots are sufficient to recover the standard model of particle physics. This approach is shown in the following. Knots and the ends of strands play no role in the model any more; the aim of highest possible simplicity is now realized.

Summary on the fundamental principle – and on continuity

We have introduced featureless, fluctuating *strands* as common constituents of space, matter, radiation and horizons. We defined fundamental events as crossing switches of strands. All physical processes are composed of fundamental events. Events and the values of all physical *observables* are defined with the help of Planck units, which in turn are due to crossing switches of strands. The definition of all physical observables through Planck units with the help of crossing switches of strands is the *fundamental principle*.

Using the fundamental principle, *continuity* of any kind – of space, fields, wave functions or time – results from the time averaging of crossing switches. This topic will be explored in detail below.

The strand conjecture fulfils the general requirements for the complete and unified
Page 149
description listed in Table 6, *provided* that it describes all motion with full precision and that it is unmodifiable.

At this point, therefore, we must start the comparison with experiment. We need to check whether strands describe *all* motion with *complete* precision. Fortunately, the task is limited: we only need to check whether strands solve each of the millennium issues
Page 164
listed in Table 8. If the strand conjecture can solve those issues, then it reproduces all observations about motion and provides a complete and unified description of nature. If the issues are not solved, the strand conjecture is worthless.

Chapter 8

QUANTUM THEORY OF MATTER DEDUCED FROM STRANDS

We show in this chapter that featureless strands that fluctuate, together with the fundamental principle – defining $\hbar/2$ as a crossing switch – imply without alternative that matter is described by quantum theory. More precisely, we deduce that tangles of fluctuating strands reproduce the spin 1/2 behaviour of matter particles, allow us to define wave functions, and imply the Dirac equation for the motion of matter. In particular, we first show that the components and phases of the wave function at a point in space are due to the orientation and phase of strand crossings at that point. Then we show that the Dirac equation follows from the belt trick (or string trick).

Furthermore, we show that strands imply the least action principle and therefore, that tangles of fluctuating strands are described by the Lagrangian of relativistic quantum particles. So far, it seems that the strand model is the only microscopic model of relativistic quantum theory that is available in the research literature.

In the present chapter, we derive the quantum theory of *matter*: we show that strands reproduce all observations about fermions and their motion. We leave for later the derivation of the quantum theory of light and the nuclear interactions, the standard model of elementary particles, and the quantum description of gravitation. As usual in quantum theory, we work in *flat* space-time.

Strands, vacuum and particles

In nature, particles move in the vacuum. The vacuum is free of matter and energy. In the strand model,

▷ ***Vacuum*** is a collection of fluctuating, unknotted and untangled strands.

Page 159 The vacuum is illustrated in Figure 14. The time average of unknotted and untangled strands has no energy and no matter content, because there are – averaged over time – *no* crossing switches and *no* tangles. The temporary crossing switches that can appear through fluctuations of the vacuum will turn out to be virtual particles; we will explore Page 239 them below. We note that the physical vacuum, being a time average, is *continuous*. The flat physical vacuum is also *unique*: it is the same for all observers. The strand model thus contains both a minimum length and a continuous vacuum. The two aspects do not contradict each other.

In nature, quantum particles *move*: quantum particles change position and phase over

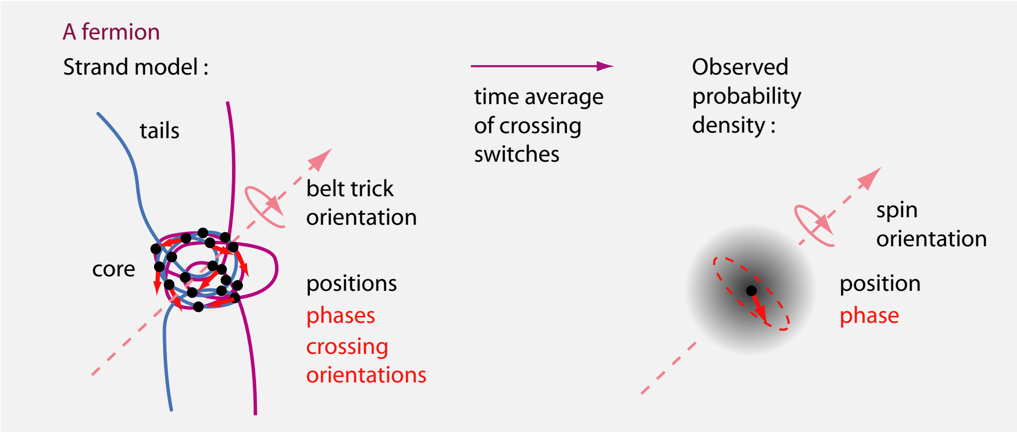


FIGURE 18 A fermion is described by a tangle of two or three strands. The crossings in the tangle core and their properties lead, after averaging, to the wave function and the probability density.

time. We therefore must define these concepts. At this stage, as just explained, we concentrate on quantum matter particles and leave radiation particles for later on. As illustrated in Figure 18 and Figure 19, we define:

▷ An elementary ***matter particle***, or ***fermion***, is a tangle of two or more strands that realizes the belt trick.

The details of this definition will become clear shortly, including the importance of the related tangle *family*. In every tangle, the important structure is the *tangle core*: the core is the that region of the tangle that contains the crossings. The core is connected to the border of space by the *tails* or *tethers* of the tangle.

▷ The ***position*** of a particle is given by the centre of the averaged tangle core. The particle position is thus the average of all its crossing positions.

▷ The ***phase*** of a matter particle is given by half the angle that describes the orientation of the tangle core around the spin axis. The particle phase is thus the average of all its crossing phases.

▷ The ***spin orientation*** of a matter particle is given by the rotation axis of the core. The spin orientation is thus the average of all its crossing orientations.

▷ The ***wave function*** of a matter particle is a blurred rendering of the crossings of its fluctuating strands.

These definitions are illustrated in Figure 18 and will be explored in detail below. We note that all these definitions imply a short-time average over tangle fluctuations. With the definitions, we get:

Page 186

▷ ***Motion*** of any quantum particle is the change of the position and orientation of its tangle core.

In nature, quantum particle motion is described by *quantum theory*. The main property of quantum theory is the appearance of the invariant quantum of action $\hbar$. In the strand model, $\hbar/2$ is described by a single crossing switch; the value of the quantum of action is thus invariant by definition.

We now explore in detail *how* the quantum of action $\hbar$ determines the motion of quantum particles. In particular, we will show that tangle fluctuations reproduce usual textbook quantum theory. As an advance summary, we clarify that

▷ ***Free quantum particle motion*** is due to fluctuations of tangle *tails*. The deformations of the tangle core are not important for free motion, and we can neglect them in this case.

In other words, when exploring quantum theory, we approximate tangle cores as being *rigid*. We will study *core deformations* in the next chapter, where we show that they are related to *interactions*. Core deformations will lead to quantum *field* theory. In this chapter we explore just the deformations of tangle *tails*; they produce the motion of *free* (and stable) quantum particles. In short, tail deformations lead to quantum mechanics. To deduce quantum mechanics from strands, we first study the rotation and then the translation of free matter particles.

Page 225

Rotation, spin 1/2 and the belt trick

In nature, quantum particles are described by their behaviour under *rotation* and by their behaviour under *exchange*. The behaviour of a particle under rotation is described by its spin value, its spin axis and its phase. The behaviour of quantum particles under exchange can be of two types: a quantum particle can be a fermion or a boson. In nature, *particles with integer spin are bosons, and particles with half-integer spin are fermions*. This is the *spin–statistics theorem*.

We now show that all properties of particle rotation and exchange follow from the strand model. We start with the case of spin 1/2 particles, and first clarify the nature of particle rotation. (We follow the usual convention to use 'spin 1/2' as a shorthand for 'z-component of spin with value $\hbar/2$'.)

It is sometimes claimed that spin is *not* due to rotation. This misleading statement is due to two arguments that are repeated so often that they are rarely questioned. First, it is said, spin 1/2 particles cannot be modelled as small rotating stones. Secondly, it is allegedly impossible to imagine rotating electric charge distributions with a speed of rotation below that of light and an electrostatic energy below the observed particle masses. These statements are correct. Despite being correct, there is a way to get around them, namely by modelling particles with strands; at the present stage, we focus on the first argument: we will show that spin *can* be modelled as rotation.

In the strand model, for all quantum particles we have:

▷ ***Spin*** is core rotation.

Indeed, in the strand model, all quantum particles, including those with spin 1/2, *differ* from everyday objects such as stones, and the essential difference is due to extension:

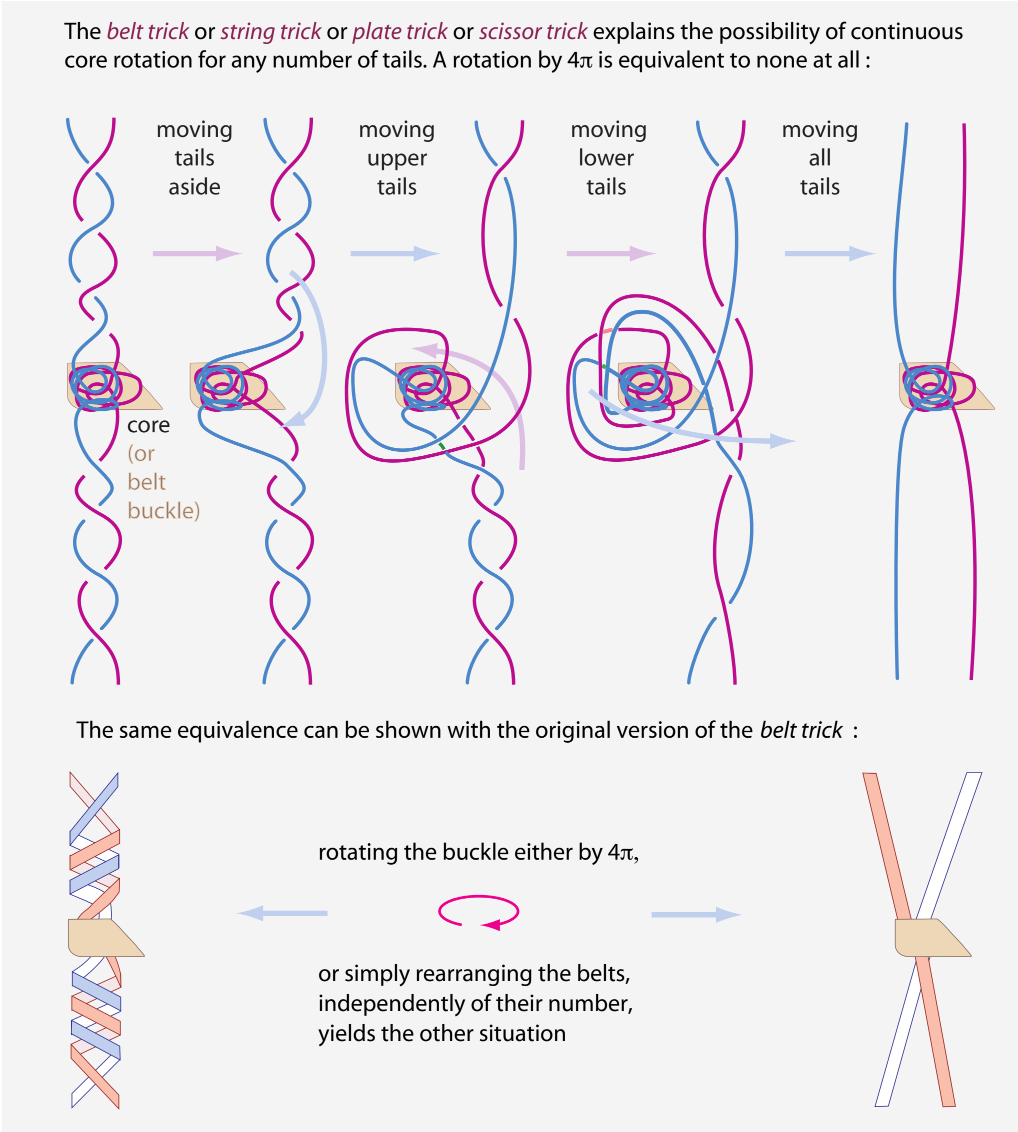


FIGURE 19 The *belt trick* – or *string trick* or *plate trick* or *scissor trick* – shows that a rotation by 4π of a central object with three or more tails (or with one or more ribbons) attached to spatial infinity is equivalent to no rotation at all. This equivalence allows a suspended object, such as a belt buckle or a tangle core, to rotate for ever. The belt trick thus shows that tangle cores made from two or more strands behave as spin 1/2 particles.

▷ ***Quantum particles*** are particles whose tails cannot be neglected.

For stones and other everyday objects, tails do not play an important role, because everyday objects are mixed states, and not eigenstates of angular momentum. In short, in everyday objects, tails can be neglected. Therefore, everyday objects are neither fermions nor bosons. But for quantum particles, the tails are essential. Step by step we will see

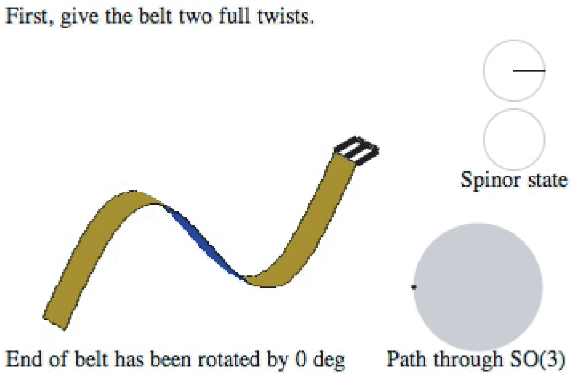


FIGURE 20 The belt trick: a double rotation of the belt buckle is equivalent to no rotation; te animation shows one way in which the belt trick can be performed. Not shown: the belt trick is also possible with *any* number of belts attached to the buckle. (QuickTime film © Greg Egan)

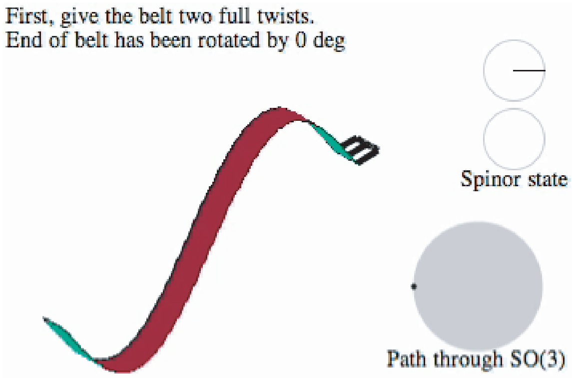


FIGURE 21 The belt trick again: this animation shows another way – another direction – in which the trick can be performed, for the same belt orientation as in the previous figure. (QuickTime film © Greg Egan)

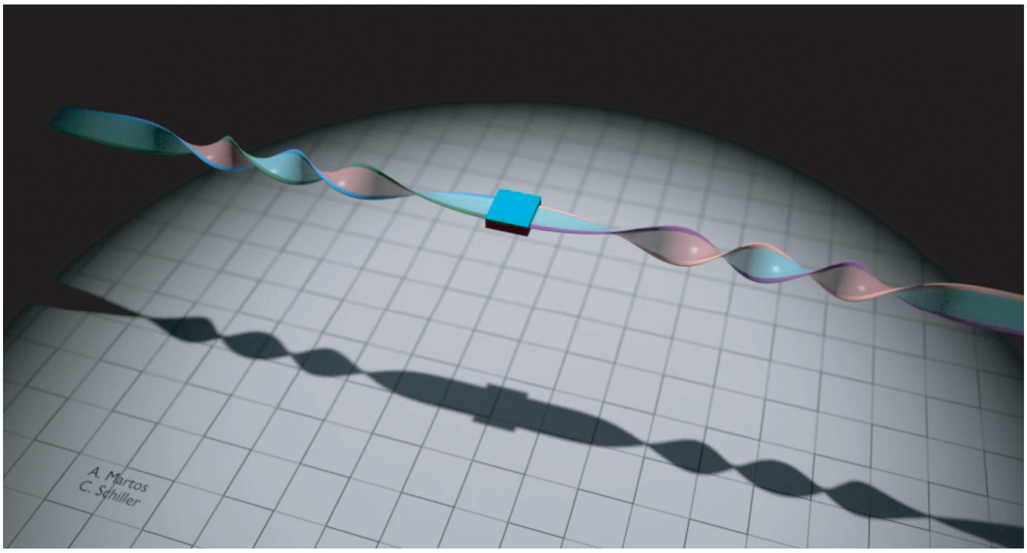
that

▷ The tails of quantum particles explain their *spin behaviour*, their *exchange behaviour* and their *wave behaviour*.

In particular, we will see that in the strand model, wave functions are *blurred* tangles; we can thus explore the general behaviour of wave functions by exploring the behaviour of tangles.

The spin behaviour of quantum particles is a consequence of strand tails. Indeed, it has been known for about a century that the so-called *belt trick* – illustrated in Figure 19, Figure 20, Figure 21 and Figure 22 – can be used, together with its variations, to model the behaviour of spin 1/2 particles under rotations. The belt trick is the observation that a belt buckle rotated by *two* full turns – in contrast to a buckle rotated by only *one* full turn – can be brought back into its original state without moving the buckle; only the motion of the belt is necessary. The belt trick is also called the *scissor trick*, the *plate trick*, the *string trick*, the *Philippine wine dance* or the *Balinese candle dance*. It is sometimes incorrectly attributed to Dirac. Ref. 162

The belt trick is of central importance in the strand model of spin 1/2 particles. In



Page 200

FIGURE 22 Assume that the belt cannot be observed, but the square object can, and that it represents a particle. The animation then shows that such a particle (the square object) can return to the starting position after rotation by 4π (but not after 2π). Such a 'belted' or 'tethered' particle thus fulfils the defining property of a *spin 1/2 particle*: rotating it by 4π is equivalent to no rotation at all. The belted square thus represents the spinor wave function; for example, a 2π rotation leads to a twist; this means a change of the sign of the wave function. A 4π rotation has no influence on the wave function. The equivalence is shown here with two attached belts, but the trick works with any positive number of belts! You can easily repeat the trick at home, with a paper strip or one or several real belts. (QuickTime film © Antonio Martos)

the strand model, all spin 1/2 particles are made of *two* (or more) tangled strands, and thus have four (or more) tails to the 'border', as shown in Figure 19. For such tangles, a rotation by 4π of the tangle core – thus a rotation by *two* full turns – can bring back the tangle to the original state, provided that the tails can fluctuate. Any system that returns to its original state after rotation by 4π is described by spin 1/2. In fact, the tails must be *unobservable* for this equivalence to hold; in the strand model, tails are simple strands and thus indeed unobservable. We will show below that the intermediate twisting of the tails that appears after rotation by only 2π corresponds to a multiplication of the wave function by -1, again as expected from a spin 1/2 particle.

Page 191

If we replace each belt by its two coloured edges, Figure 22 shows how tails behave when a spin 1/2 tangle is rotated. By the way, systems with tails – be they strands or bands – are the *only possible systems* that realize the spin 1/2 property. Only systems with tails to spatial infinity have the property that a rotation by 4π is equivalent to no rotation at all. (Can you show this?) The fundamental connection between spin 1/2 and extension is one of the properties that led to the strand model.

Challenge 127 e

The animations show that the belt trick works with one and with two belts attached to the buckle. In fact, belt trick works with *any number* of belts attached to the buckle. The belt trick even works with *infinitely* many belts, and also with a *full two-dimensional sheet*. The wonderful video www.youtube.com/watch?v=UtdljdoFAwg by Gareth Taylor

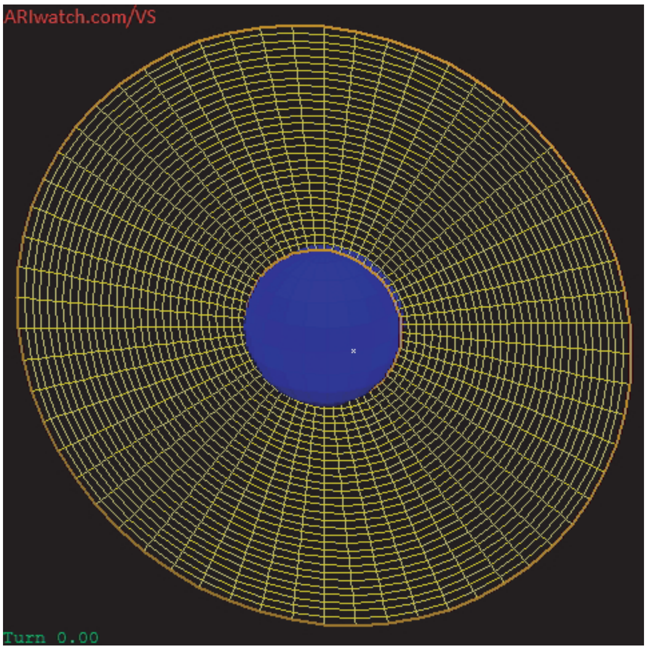


FIGURE 23 The belt trick realized as a rotating ball attached to a sheet (QuickTime film © www.ariwatch.com/VS/Algorithms/DiracStringTrick.htm).

and the slightly different animation of Figure 23 both illustrate the situation. A sphere glued to a flexible sheet can be rotated as often as you want: if you do this correctly, there is *no* tangling and you can go on for ever.

Vol. I, page 168 In fact, the 'belt' trick also works if the central object to be rotated is glued into a flexible, three-dimensional jelly or mattress. This was shown early on in our adventure.

The animations of the belt trick lead us to a statement on strands and tangles that is central for the strand model:

▷ An object or a tangle core that is attached by (three or more) tails to the border of space can rotate *continuously*.

Here we made the step from belts to strands. In other terms, the possibility of continuous rotation allows us to describe spin 1/2 particles by rotating tangles. In other terms,

▷ Rotating tangles model spin.

The tail fluctuations required to rearrange the tails after two full turns of the core can be seen to model the average *precession* of the spin axis. We thus confirm that spin and rotation are the same for spin 1/2 particles.

The belt trick is not unique

Ref. 162 One aspect of the belt trick seems unmentioned in the research literature: after a rotation of the belt buckle or tangle core by 4π, there are various options to untangle the tails. Two

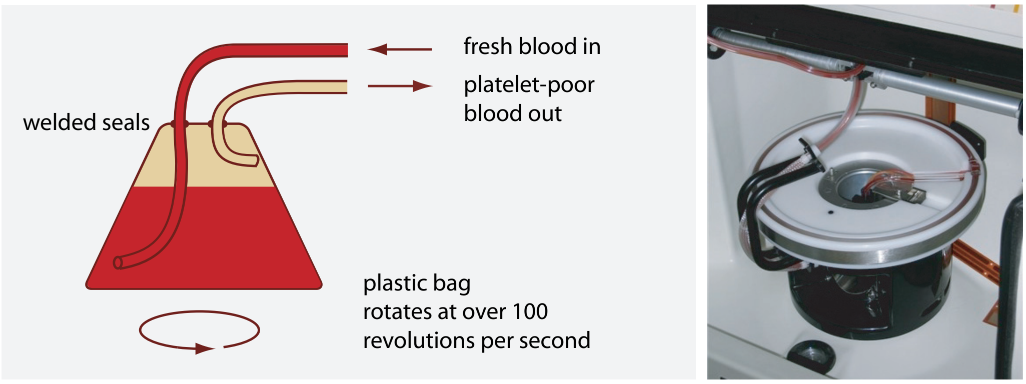


FIGURE 24 In an apheresis machine, the central bag spins at high speed despite being connected with tubes to a patient; this happens with a mechanism that continuously realizes the belt trick (photo © Wikimedia).

different options are shown in Figure 20 and Figure 21. You can test this yourself, using a real belt. In fact, there are *two* extreme ways to perform the belt trick, and a continuum of options in between. These options will be of central importance later on: the options require a description of fermions with *four* complex functions. We will discover that the various options of the belt trick are related to the difference between matter and antimatter and to the parity violation of the weak interaction.

Challenge 128 e

An aside: the belt trick saves lives

Without the belt trick, the *apheresis machines* found in many hospitals would not work. When a person donates blood platelets, the blood is continuously extracted from one arm and fed into a bag in a centrifuge, where the platelets are retained. The platelet-free blood then flows back into the other arm of the donor. This happens continuously, for about an hour or two. In order to be *sterile*, tubes and bag are used only once and are effectively *one* piece, as shown in Figure 24. Apheresis machines need *tethered rotation* to work. Topologically, this set-up is identical to a fermion tangle: each tube corresponds to one belt, or two strand tails, and the rotating bag corresponds to the rotating core.

In such apheresis machines, the centrifugation of the central bag takes place at over 100 revolutions per second, in the way illustrated in Figure 25. To avoid tangling up the blood tubes, a bracket moves the tubes during each rotation, alternatively up and down. This so-called *anti-twister mechanism* produces precisely the motion along which the belt moves when it is untangled after the buckle is rotated by 4π. An apheresis machine thus performs the belt trick 50 times per second, with each rotation of the centrifugation. Due to the centrifugation, the lighter platelets are retained in the bag, and the heavier components of the blood are pumped back to the donor. The retained platelets are then used to treat patients with leukaemia or severe blood loss due to injury. A single platelet donation can sustain several lives.

In short, without the belt trick, platelet donations would not be sterile and would thus be impossible. Only the belt trick, or tethered rotation, allows sterile platelet donations that save other people's lives.

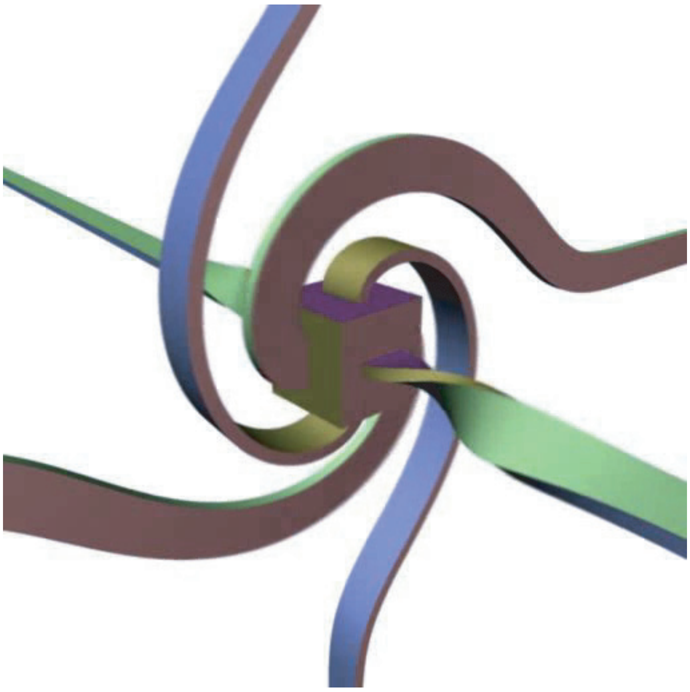


FIGURE 25 The basis of the apheresis machine – and yet another visualisation of the belt trick, here with 6 belts (QuickTime film © Jason Hise).

FIGURE 26 A version of the antitwister mechanism, or belt trick, with 96 belts attached to a black and white ball that rotates continuously (QuickTime film © Jason Hise).

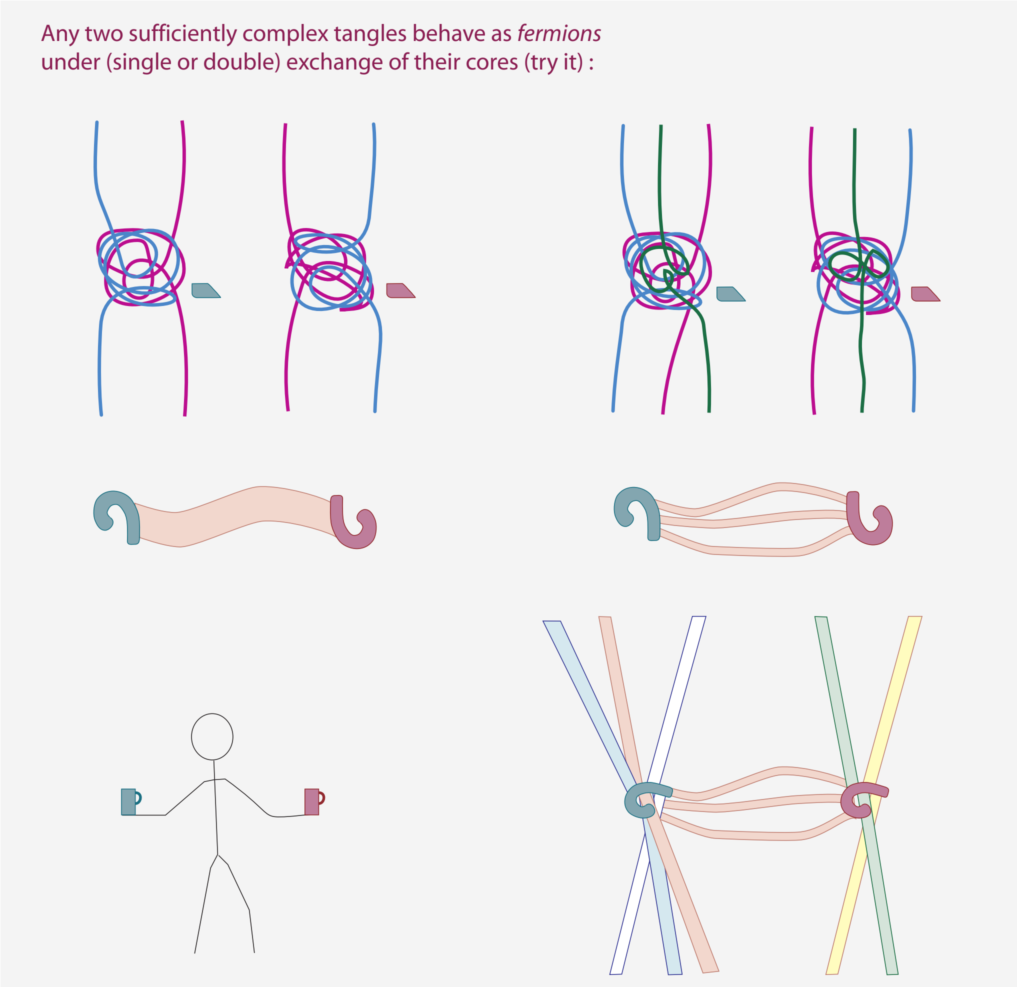


FIGURE 27 When two spin 1/2 tangles each made of several strands or bands, are exchanged *twice*, it is possible to rearrange their tails to yield the original situation. This is not possible when the tangles are only exchanged *once*. Spin 1/2 tangles are thus fermions. The figure presents common systems that show this behaviour.

Fermions and spin

In nature, *fermions* are defined as those particles whose wave function changes sign when they are exchanged. Does the strand model reproduce this observation?

We will see below that in the strand model, wave functions are *blurred* tangles. We thus can explore exchange properties of quantum particles and of their wave functions by exploring the exchange properties of their tangles. Now, if we exchange two tangle cores *twice*, while keeping all tails connections fixed, tail fluctuations alone can return the situation back to the original state! The exchange properties of spin 1/2 tangles are easily checked by playing around with some pieces of rope or bands, as shown in Figure 27, or by watching the animation of Figure 28.

The simplest possible version of the experiment is the following: take two coffee cups,

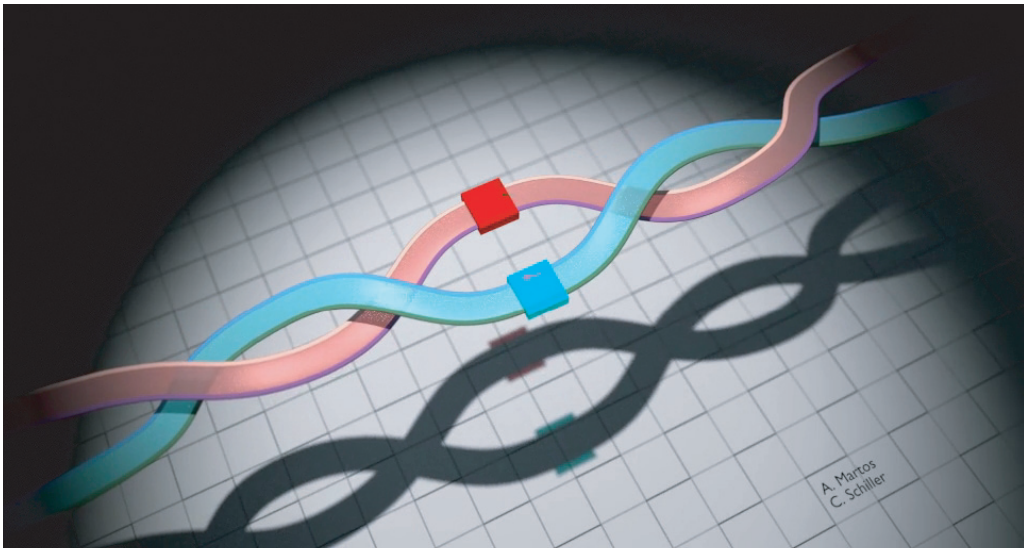


FIGURE 28 Assume that the belts cannot be observed, but the square objects can, and that they represent particles. We know from above that belted buckles behave as spin 1/2 particles. The animation shows that two such particles return to the original situation if they are switched in position twice (but not once). Such particles thus fulfil the defining property of *fermions*. (For the opposite case, that of bosons, a simple exchange would lead to the identical situation.) You can repeat the trick at home using paper strips. The equivalence is shown here with two belts per particle, but the trick works with any positive number of belts attached to each buckle. This animation is the essential part of the proof that spin 1/2 particles are fermions. This is called the *spin–statistics theorem*. (QuickTime film © Antonio Martos)

one in each hand, and cross the two arms over each other (once). Keeping the orientation of the cups fixed in space, uncross the arms by walking around the cups. This is possible, but as a result, both arms are twisted. If you are intrepid, you can repeat this with two (or more) people holding the cups. And you can check the difference with what is possible after a *double* crossing of arms: in this case, all arms return to the starting situation.

All these experiments show:

▷ A *simple* exchange of two spin 1/2 particles (tangles, cups on hands, belt buckles) is equivalent to a multiplication by -1, i.e., to *twisted* tangles, arms or belts.

▷ In contrast, a *double* exchange of two spin 1/2 particles can always be untwisted and is equivalent to no exchange at all.

Spin 1/2 particles are thus fermions. In other words, the strand model reproduces the spin–statistics theorem for spin 1/2: all elementary matter particles are fermions. In summary, a tangle core made of two or more tangled strands behaves – both under rotations and under exchange – like a spin 1/2 particle.

We note that it is sometimes claimed that the appearance of spin 1/2 can only be mod-

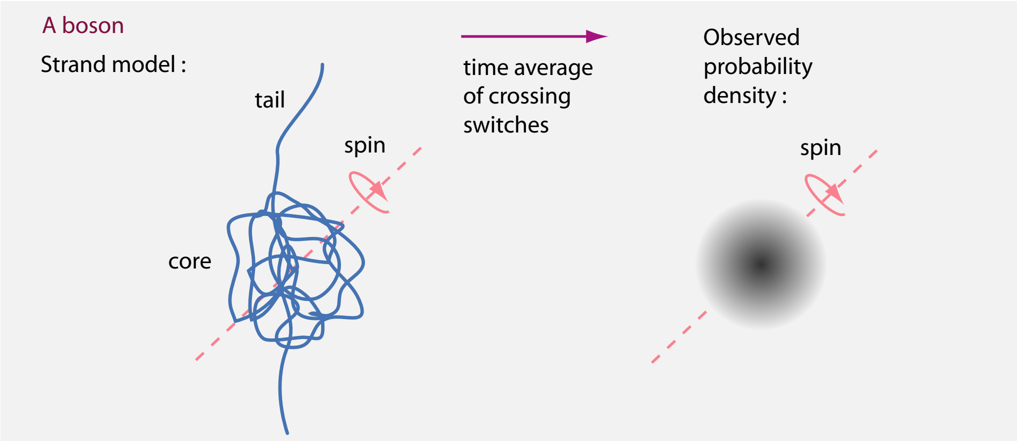


FIGURE 29 A massive spin 1 particle in the strand model (left) and the observed probability density when averaging its crossings over long time scales (right).

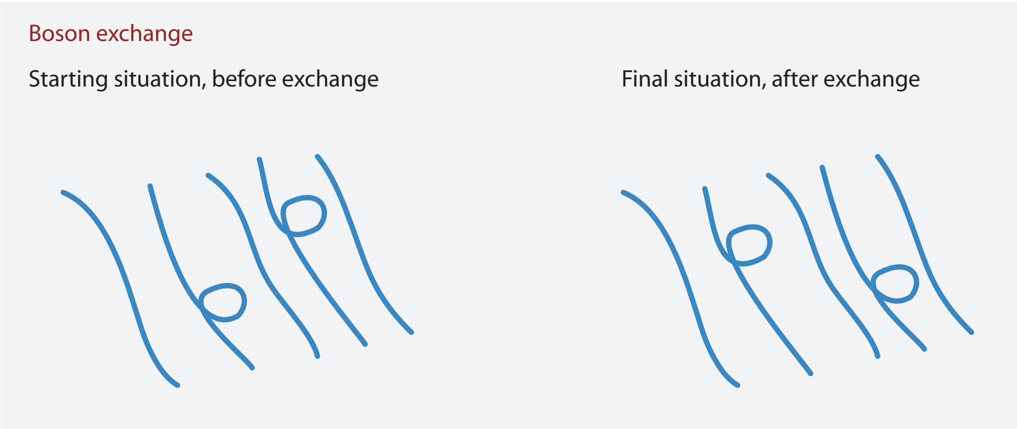


FIGURE 30 In the strand model, unknotted *boson* tangles can switch states without generating crossings, and thus without changing the sign of the phase.

elled with the help of a topology change of space or space-time. The various belt trick animations given above prove that this is not correct: spin 1/2 can be modelled in three dimensions in all its aspects. No topology change is needed. You might want to model the creation of a spin 1/2 particle–antiparticle pair as a final argument. Challenge 129 e

Bosons and spin

For tangles made of *one* strand – thus with two tails to the border – a rotation of the tangle core by 2π restores the original state. Such a tangle, shown in Figure 29, thus behaves like a spin 1 particle. The figure also shows the wave function that results from time averaging the crossings.

Bosons are particles whose combined state does not change phase when two particles are exchanged. We note directly that this is impossible with the tangle shown in Figure 29; the feat is only possible if the boson tangle is made of *unknotted* strands. Indeed, Vol. IV, page 117

for unknotted strands, the exchange process can easily switch the two deformations, as illustrated in Figure 30.

▷ Massive elementary particles thus can only be bosons if they also have an *unknotted* tangle in the tangle family that represents them.

The simplest strand model for each elementary boson – the photon, the W boson, the Z boson, the gluon and the Higgs boson – must thus be made of unknotted strands. We Page 314, page 330 will deduce the precise tangles below, in the chapter on the particle spectrum. The tangle for the hypothetical graviton – also a boson, but in this case with spin 2 and invariant Page 281 under core rotations by π – will be introduced in the chapter on general relativity.

In summary, unknotted tangles realize the spin–statistics theorem for particles with integer spin: radiation particles, which have integer spin, are automatically bosons.

Spin and statistics

We just saw that fluctuating strands reproduce the spin–statistics theorem for fermions and for bosons, and thus for all elementary particles, if appropriate tangles are used. Apart from this fundamental result, the strand model also implies that no spins lower Challenge 130 e than $\hbar/2$ are possible, and that spin values are always an *integer multiple of* $\hbar/2$. All this matches observations.

In the strand model, temporal evolution and particle reactions *conserve* spin, because all interactions conserve the number of strands and tails. The details of the conservation will become clear later on. Again, the result agrees with observations.

The strand model thus explains the origin of permutation symmetry in nature: *permutation symmetry* of particles is due the possibility to exchange tangle cores of identical particles; and identical particles have tangle cores of identical topology. We have thus Page 164 already ticked off one item from the millennium list of unexplained properties of nature.

In summary, the strand model reproduces the rotation, the spin and the exchange behaviour of elementary quantum particles – both fermions and bosons – in all its observed details. We now proceed to the next step: quantum mechanics of translational motion.

Tangle functions: blurred tangles

In the strand model, particle motion is due to the motion of tangle cores. But according to the fundamental principle, strands and tangles are not observable; only crossing switches are. To explore the relation between crossing switches and motion, we first recall what a crossing is.

▷ A ***crossing*** of strands is a local minimum of strand distance. The position, orientation and the phase of a crossing are defined by the space vector corresponding to the local distance minimum, as shown in Figure 31. The sign of the orientation is defined by arbitrarily selecting one strand as the starting strand. The even larger arbitrariness in the definition of the phase will be of great importance later on, and lead to gauge invariance.

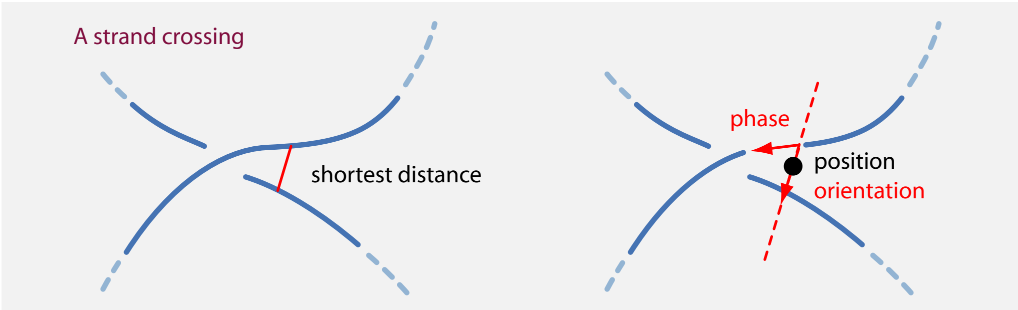


FIGURE 31 The definition of a crossing, its position, its orientation and its phase.

To describe the motion of tangles, we need concepts that allow us to take the step from general strand fluctuations to the motion of tangle cores. As a mathematical tool to describe crossing fluctuations, we define:

▷ The ***tangle function*** of a system described by a tangle is the short-time average of the positions and the orientations of its crossings (and thus *not* of crossing switches and *not* of the strands themselves).

The tangle function can be called the 'oriented crossing density' or simply the 'blurred tangle'. As such, the tangle function is a continuous function of space, similar to a cloud; we will see below what its precise mathematical description looks like. The tangle function captures the short-time average of all possible tangle fluctuations. For a tangle made of two strands, Figure 32 illustrates the idea. However, the right-hand side of the figure does not show the tangle function itself, but its probability density. We will see shortly that the probability density is the (square of the) crossing *position* density, whereas the tangle function is a density that describes *both position and orientation* of crossings.

The tangle function at any given time is *not* observable, as its definition is not based on crossing switches, but only on crossings. However, since crossing switches only occur at places with crossings, the tangle function is a useful tool to *calculate* observables. In fact, we will show that the tangle function is just another name for what is usually called the *wave function*. In short, the tangle function, i.e., the oriented crossing density, will turn out to describe the *quantum state* of a system.

In summary, the tangle function is a blurred image of the tangle – with the important detail that the crossings are blurred, not the strands.

▷ For the definition of the tangle function, the ***short-time average*** of crossings is taken over the typical time resolution of the observer. This is a time that is much *longer* than the Planck time, but also much *shorter* than the typical evolution time of the system. The time resolution is thus what the observer calls an 'instant' of time. Typically – and in all known experiments – this will be 10^{-25} s or more; the typical averaging will thus be over a time interval with a value between 10^{-43} s, the Planck time, and around 10^{-25} s.

There are *two ways* to imagine tangle fluctuations and to deduce the short-time average

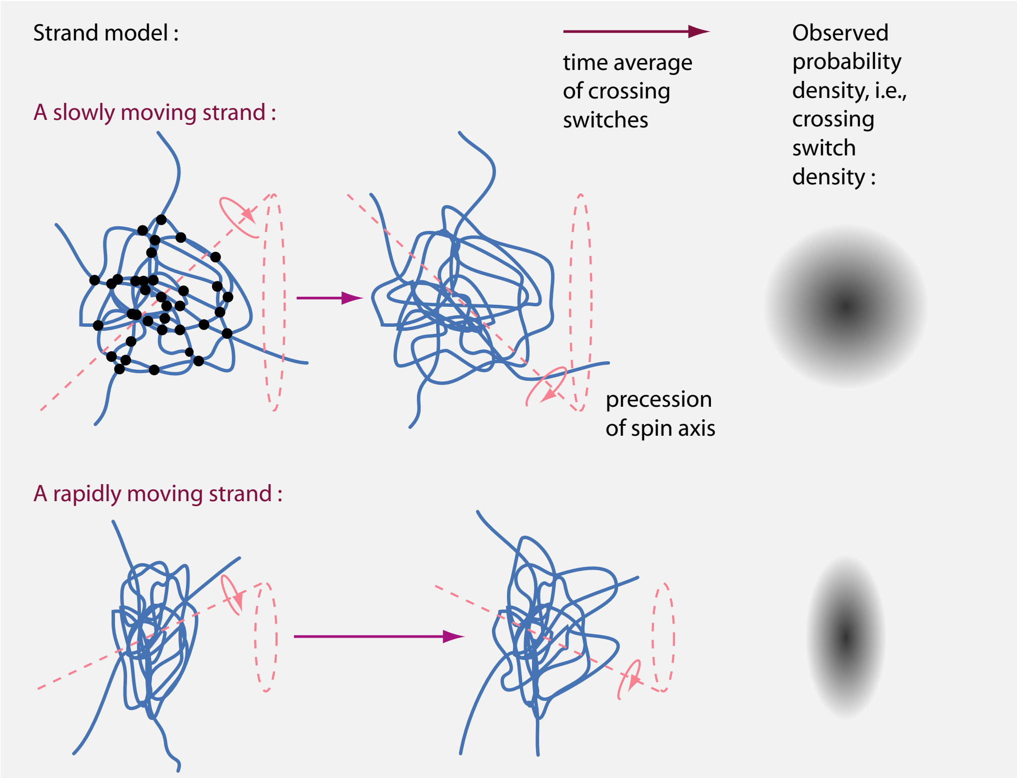


FIGURE 32 Some strand configurations, some of their short time fluctuations, and the corresponding probability density that results when averaging crossing switches over time. (The black dots are not completely drawn correctly.)

from a given tangle. The first, straightforward way is to average over all possible strand fluctuations during the short time. *Each piece of strand* can change in shape, and as a result, we get a cloud. This is the common *Schrödinger picture* of the wave function and of quantum mechanics. The second, alternative way to average is to imagine that the *tangle core as a whole* changes position and orientation randomly. This is easiest if the core with all its crossings is imagined to be tightened to a small, almost 'point-like' region. Then all observables are also localized in that region. It is often simpler to imagine an average over all position and orientation fluctuations of such a tightened core, that to imagine an average over all possible strand fluctuations. This alternate view leads to what physicists call the *path integral formulation* of quantum mechanics. (Can you show the equivalence of the two averaging methods?) Of course, in both cases the final result is that the tangle function is a cloud, i.e., a probability amplitude.

Challenge 131 e

Details on fluctuations and averages

In the strand model, the strand fluctuations of particle strands are a consequence of the embedding of all particles in a background which itself is made of fluctuating vacuum strands. Fluctuations randomly add detours to particle strands and randomly shift the core position. Fluctuations do not keep the strand length constant. Fluctuations do not conserve strand shape nor any other property of strands, as there is no mechanism that enforces such rules. Strand fluctuations are thus quite wild. What then can be said about the details of the averaging procedure for strand fluctuations?

The *fluctuations of the vacuum* are those strand fluctuations that lead to the definition of the background space. This definition is possible in a consistent manner only if the fluctuations are homogeneous and isotropic. The vacuum state can thus be *defined* as that state for which the fluctuations are (locally) homogeneous and isotropic. In particular, the fluctuations imply

▷ Flat vacuum has a tangle function that vanishes everywhere.

Challenge 132 e

The proof is an interesting exercise. The existence of a homogeneous and isotropic background space then implies conservation of energy, linear and angular momentum of particles travelling through it.

The *fluctuations of a tangle* lead, after averaging, to the tangle function, i.e., as we will see, to the wave function. The conservation of energy and momentum implies that the time average of the tangle fluctuations also conserves these quantities.

Page 296
Page 110, page 169

Therefore we can continue our discussion without yet knowing the precise details of the tangle fluctuations themselves. (We will provide these details below, in the section on general relativity.) Here we only require that the average of the fluctuations behaves in such a way as to be *consistent* with the definition of the background used by the observer. We thus make explicit use of the conviction that a background-free description of nature is impossible, and that a fundamental description of nature *must* contain a circular definition that makes an axiomatic description of nature impossible. Despite this limitation, such a circular description of nature must be *self-consistent*.

Page 205

We will also show below that the definition of the tangle function does *not* introduce hidden variables, even though first impression might suggest the opposite. In fact, it is possible to define something akin to a strand evolution equation. However, it does not deepen our understanding of the evolution equation of the wave function.

Tangle functions are wave functions

In the following, we show that the *tangle function*, the blurred image of tangle crossings, is the same as what is usually called the wave function. We recall what we know from textbook quantum theory:

▷ A single-particle ***wave function*** is, generally speaking, a *rotating and diffusing cloud*.

The *rotation* describes the evolution of the phase, and the *diffusion* describes the evolution of the density. We now show that tangle functions have these and all other known properties of wave functions. We proceed by deducing all the properties from the definition of tangle functions. We recall that, being a short-time average, a tangle function is a continuous function of space and time.

▷ Using the tangle function, we define the strand ***crossing position density***, or ***crossing density***, for each point in space, by discarding the orientation information, counting the crossings in a volume, and taking the square root. The crossing density – more precisely, its square root – is a *positive number*, more precisely, a positive real function $R(x,t)$ of space and time.

We will see shortly that the crossing position density is the square root of what is usually called the *probability density*.

▷ A tangle function also defines an ***average crossing orientation*** and a ***average phase*** at each point in space. The average crossing orientation and the average phase are related to the *spin orientation* and *phase* of the wave function. The mathematical descriptions of these quantities depend on the approximation used.

The *simplest approximation* for a tangle function is to assume, in the physical situation under study, that the spin direction is *independent* of spatial position and thus not taken into consideration; this approximation will lead to the Schrödinger equation. In this simplest approximation, at each point in space, the local average orientation of the fluctuations of the tangle core will just be described by a *single angle*. This quantum phase is a function of time and space and describes how much the local average phase is rotated around the fixed spin orientation.

▷ The ***quantum phase*** of fermions is *one half* the core rotation angle α.

Without the neglect of spin, and especially when the spin axis can *change* over space, the description of orientation and phase averages require more details; we will study these cases separately below. They will lead to the non-relativistic Pauli equation and to the relativistic Dirac equation.

In short, in the simple approximation when spin effects can be neglected, the local tangle function value can be described by one real number R and by one quantum phase α. The tangle function can thus be described by a *complex number* ψ at each point in space and time:

$$\psi(x,t) = R(x,t)\, e^{i\alpha(x,t)/2} \; . \tag{122}$$

If a system changes with time, the tangle function changes; this leads to crossing switches; therefore, temporal evolution is expected to be observable through these crossing switches. As we will see shortly, this leads to an evolution equation for tangle functions.

Challenge 133 s Here is a fun challenge: how is the shortest distance between the strands, for a crossing located at position x and t, related to the magnitude, i.e., the absolute value $R(x,t)$, of the wave function?

Page 206 We note that if *many* particles need to be described, the many-particle tangle function defines a separate crossing density for each particle tangle.

Tangle functions form a *vector space*. To show this, we need to define the ***linear combination*** or ***superposition*** $\chi = a_1\psi_1 + a_2\psi_2$ of two tangle functions. This requires the definition of two operations: scalar multiplication and addition. We can do this in two ways. The first way is to define the operations for tangle functions directly, as is done in quantum mechanics:

▷ First, boring definition: The ***scalar multiplication*** $a\psi$ and the ***addition*** $\psi_1 + \psi_2$ of quantum states are taken by applying the relative operations on complex numbers at each point in space, i.e., on the local values of the tangle function.

The second way to deduce the vector space is more fun, because it will help us to visualize quantum mechanics: we can define addition and multiplication for tangles, and take the time average *after* the tangle operation is performed.

▷ Second, fun definition: The ***scalar multiplication*** $a\psi$ of a state ψ by a complex number $a = re^{i\delta}$ is formed by taking a tangle underlying the tangle function ψ, then rotating the tangle core by the angle 2δ, and finally pushing a fraction $1-r$ of the tangle to the border of space, thus keeping the fraction r of the original tangle at finite distances. Time averaging then leads to the tangle function $a\psi$.

The scalar multiplication for strands is illustrated in Figure 33. The above definition of scalar multiplication is only defined for factors $r \leqslant 1$. Indeed, no other factors ever appear in physical problems (provided all wave functions are normalized), so that scalar multiplication is not required for other scalars.

The strand version of scalar multiplication is *unique*; indeed, even though there is a choice about which fraction r of a tangle is kept and which fraction $1 - r$ is sent to the border of space, the resulting tangle function, which is defined as an average over fluctuations, is independent from this choice.

The scalar multiplication of strands behaves as expected for 1 and 0. By construction, the strand version of scalar multiplication is associative: we have $a(b\psi) = (ab)\psi$. The strand multiplication by -1 is defined as the rotation of the full tangle core by 2π.

We also need to define the addition operation that appears in the linear combination of two tangle functions. This is a straightforward complex addition at each point in space. Again, for fun, we also define the operation on tangles themselves, and take the time average that leads to the tangle function afterwards.

▷ Second, fun definition: The ***addition*** of two tangles $a_1\psi_1$ and $a_2\psi_2$, where ψ_1 and ψ_2 have the same topology and where $a_1^2 + a_2^2 = 1$, is defined by connecting those tails that reach the border of space, and discarding all parts of the tangles that were pushed to the border of space. The connection of tangles must be performed in such a way as to maintain the topology of the original tangles; in particular, the connection must not introduce any crossings or linking. Time averaging then leads to the tangle function of the superposition $\chi = a_1\psi_1 + a_2\psi_2$.

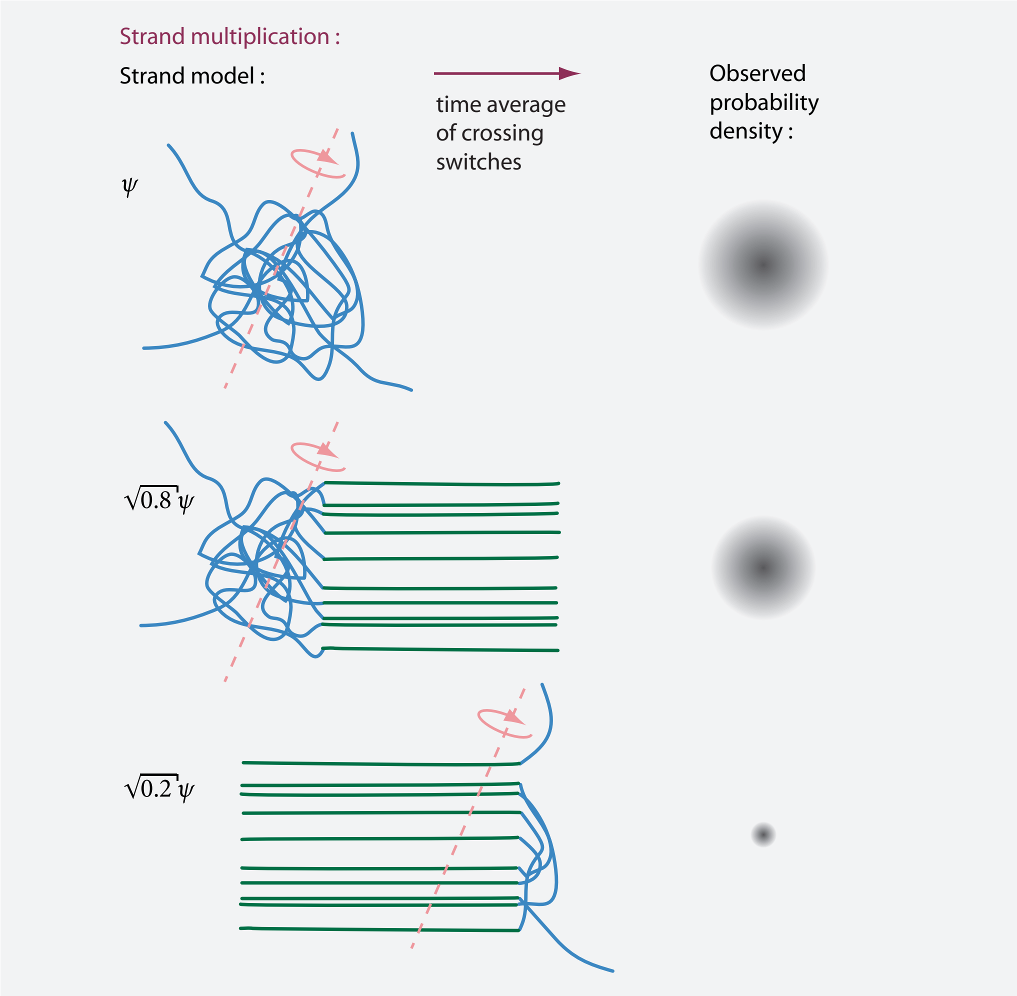


FIGURE 33 Scalar multiplication of localized tangles, visualizing the scalar multiplication of wave functions.

To visualize the result of addition and superposition, it is easiest to imagine that the strands reaching the border of space have fluctuated back to finite distances. This is possible because by definition, these connections are all unlinked. An example of superposition, for the case of two quantum states at different positions in space, is shown in Figure 34. We note that despite the wording of the definition, no strand is actually cut or re-glued in the operation of addition.

Page 193 The definition of linear combination requires that the final strand χ has the same topology and the same norm as each of the two strands ψ_1 and ψ_2 to be combined. Physically, this means that only states for the same particle can be added and that particle number is preserved; this automatically implements the so-called *superselection rules* of quantum theory. This result is pretty because in usual quantum mechanics the superselection rules need to be added by hand. This is not necessary in the strand model.

The sum of two tangle functions is *unique*, for the same reasons given in the case of

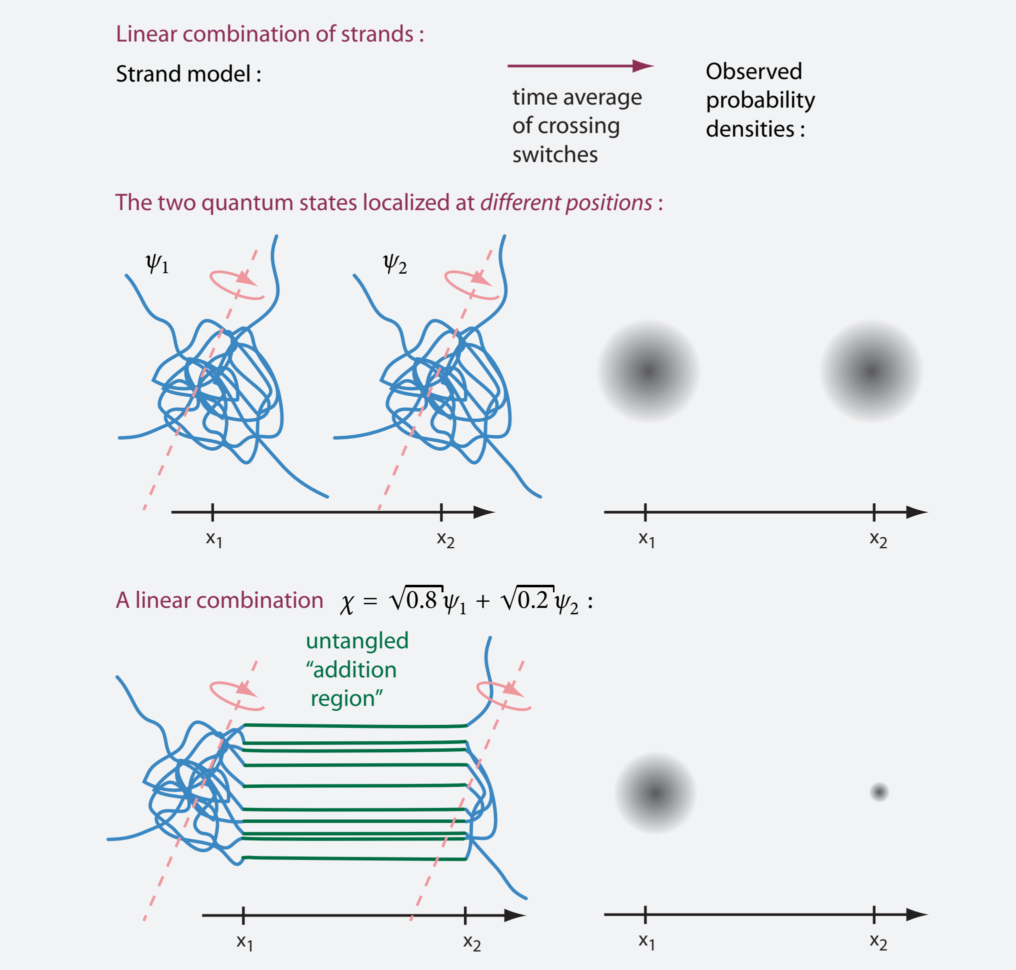


FIGURE 34 A linear combination of strands, in this case for two states representing a particle at two different position in space, visualizing the linear combination of wave functions.

scalar multiplication. The definition of addition can also be extended to more than two terms. Addition is commutative and associative, and there is a zero state, or identity element, given by no strands at all. The definition of addition also implies distributivity with respect to addition of states and with respect to addition of scalars. It is also possible to extend the definitions of scalar multiplication and of addition to all complex numbers and to unnormed states, but this leads us too far from our story.

Challenge 134 e

In short, tangle functions form a vector space. We now define the scalar product and the probability density in the same way as for wave functions.

▷ The ***scalar product*** between two states φ and ψ is $\langle\varphi|\psi\rangle = \int \overline{\varphi}(\mathbf{x})\psi(\mathbf{x})\,\mathrm{d}\mathbf{x}$.

▷ The ***norm*** of a state is $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$.

▷ The ***probability density*** ρ is $\rho(x,t) = \overline{\psi}(x,t)\psi(x,t) = R^2(x,t)$. It thus ignores the orientation of the crossings and is the *crossing position density*.

The scalar product and the probability density are *observables*, because their definitions can be interpreted in terms of crossing switches. Indeed, the scalar product $\langle\varphi|\psi\rangle$ can be seen as the (suitably normed) number of crossing switches required to transform the tangle $\overline{\varphi}$ into the tangle ψ, where the tangle $\overline{\varphi}$ is formed from the tangle φ by exchanging the orientation of each crossing. A similar interpretation is possible for the probability density, which therefore is at the same time the crossing density squared and the crossing switch density. We leave this confirmation as fun for the reader.

Challenge 135 e

It is also possible to define the scalar product, the norm and the probability density using tangles, instead of using tangle functions. This is left as a puzzle to the reader.

Challenge 136 e

In summary, we have shown that tangle functions form a *Hilbert space*. The next steps are now obvious: We must first show that tangle functions obey the Schrödinger equation. Then we must extend the definition of quantum states by including spin and special relativity, and show that they obey the Dirac equation.

Deducing the Schrödinger equation from tangles

The Schrödinger equation, like all evolution equations in the quantum domain, results when the definition of the wave function is combined with the energy–momentum relation. As already mentioned, the Schrödinger equation for a quantum particle also assumes that the orientation of particle spin is constant for all positions and all times. In this case, the spin can be neglected, and the tangle function is a single complex number at each point in space and in time, usually written $\psi(x, t)$. How does the tangle function evolve in time? To answer this question, we will only need the fundamental principle that crossing switches define the quantum of action $\hbar$.

We start with a free particle. We assume a fixed, but *unspecified* rotation direction of its tangle. Now, in the strand model, a localized particle with constant speed is described by a localized tangle that rotates and advances. In other words, the strand fluctuations produce a peak of probability density that changes position with constant speed.

Every tangle rotation leads to crossing switches. A rapid tangle rotation leads to many crossing switches per time, and slow rotation to few crossing switches per time. Now, the fundamental principle tells us that crossing switches per time are naturally measured in action per time, or ***energy***. In other words, tangle rotation is related to tangle energy.

- ▷ Particles with *high* energy have *rapidly* rotating tangles.
- ▷ Particles with *low* energy have *slowly* rotating tangles.

*The **energy** of a rotating tangle is the number of crossing switches per time.* Rotating a tangle core leads to crossing switches in its tails. In the strand model, the kinetic energy E of a particle is thus due to the crossing switches formed in its tails. In other words, the kinetic energy E is related to the (effective) angular frequency ω of the core rotation by

$$E = \hbar\omega \, . \tag{123}$$

The local phase of the tangle function ψ changes with the rotation. This implies that

$$\omega = i\partial_t \psi \, . \tag{124}$$

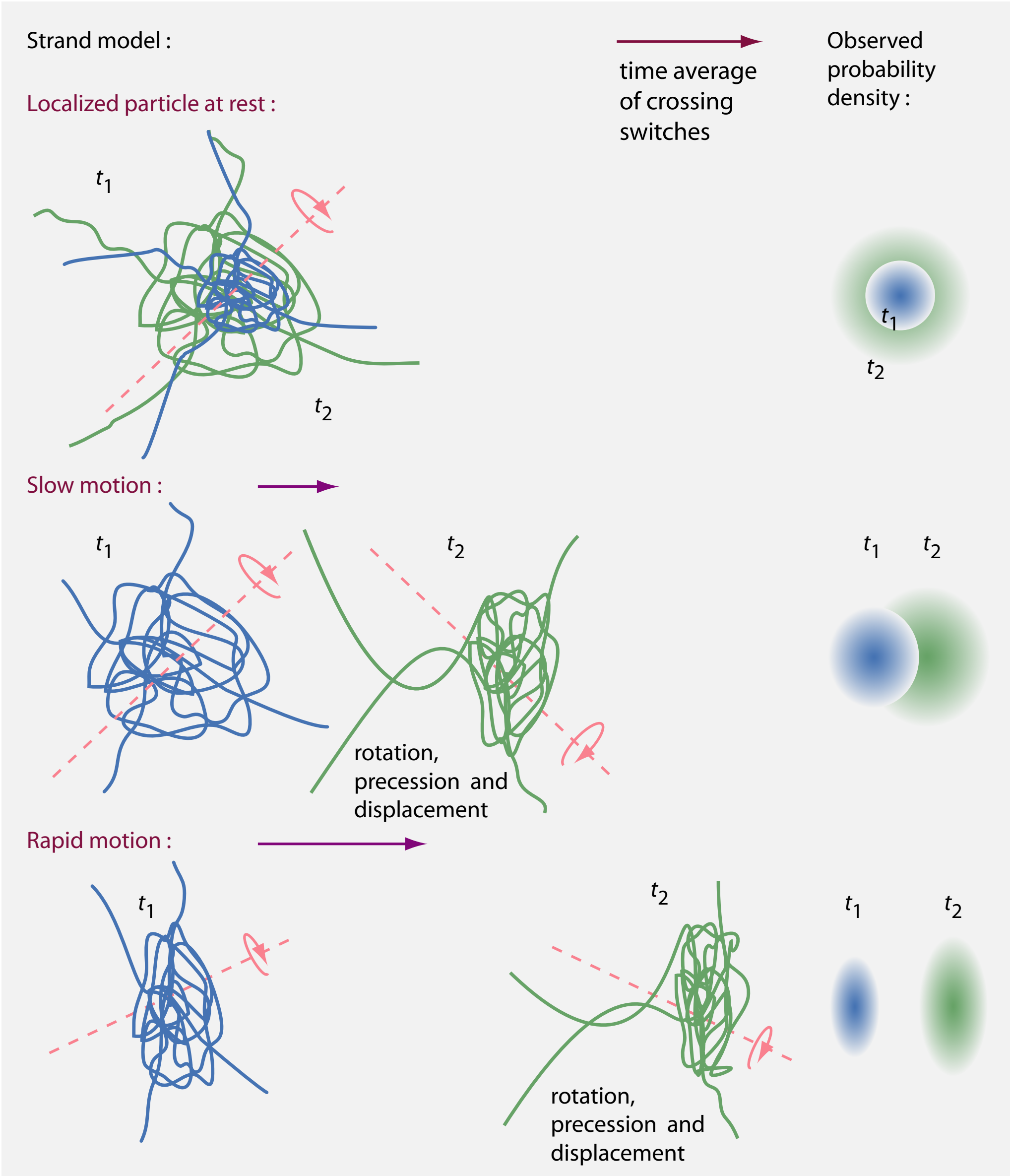


FIGURE 35 Examples of moving tangles of free particles.

We will need the relation shortly.

The linear motion of a tangle implies that it makes also sense to pay attention to the number of crossing switches per distance.

▷ Rapidly moving tangles show many crossing switches per distance.

▷ Slowly moving tangles show few crossing switches per distance.

The fundamental principle tells us that the natural observable to measure crossing switches per distance is action per distance, or ***momentum***. Linear motion of tangles is thus related to momentum: *The **momentum** of a moving tangle is the number of crossing switches per distance.* The momentum p is thus related to the (effective) wave number $k = 2\pi/\lambda$ of the core motion by

$$p = \hbar k . \tag{125}$$

The local phase of the tangle function ψ changes with the motion. This implies

$$k = -i\partial_x \psi . \tag{126}$$

This completes the description of matter wave functions without spin.

The belt trick for the fluctuating tails now has a fascinating consequence. To allow the belt trick also for high linear momentum, the more the momentum increases, the more the spin rotation axis has to align with the direction of motion. This is shown in Figure 35. This leads to a quadratic increase of crossing switches with momentum p: one factor p is due to the increase of the speed of rotation, the other factor is due to the increase of the alignment. We thus get

$$E = \frac{p^2}{2m} \quad \text{and} \quad \omega = \frac{\hbar}{2m} k^2 . \tag{127}$$

This is dispersion relation for masses moving at velocities much smaller than the speed of light. The relation agrees with all experiments. The constant m is a proportionality factor that depends on the tangle core. We can now use the same argument that was used already by Schrödinger. Substituting the tangle relations in the dispersion relation, we get the evolution equation for the tangle function ψ given by

$$i\hbar\partial_t \psi = -\frac{\hbar^2}{2m} \partial_{xx} \psi . \tag{128}$$

This is the famous Schrödinger equation for a free particle (written for just one space dimension for simplicity). We thus have deduced the equation from the strand model under the condition that spin can be neglected and that velocities are small compared to the speed of light. In this way, we have also deduced, indirectly, Heisenberg's *indeterminacy relations*.

We have thus completed the proof that tangle functions, in the case of negligible spin effects and small velocities, are indeed wave functions. In fact, tangle functions are wave functions also in the more general case, but then their mathematical description is more involved, as we will see shortly. We can sum up the situation in a few simple terms: *wave functions are blurred tangles.*

Mass from tangles

In quantum theory, particles spin while moving: the quantum phase rotates while a particle advances. The coupling between rotation and translation has a name: it is called the *mass* of a particle. We saw that the rotation is described by an average angular frequency ω, and the translational motion is described by a wave number k. The proportionality factor $m = \hbar k^2/2\omega = p^2/2E$ is thus a quantity that relates rotation frequency and wave number. In quantum theory,

▷ The (inertial) mass m describes the coupling between translation and rotation.

We note that a large mass value implies, for a given momentum value, both a slow translation and a slow rotation.

In the strand model, particle translation and rotation are modelled by the translation and rotation of the tangle core. Now, the strand model makes a point that goes beyond usual quantum theory. The strand model explains *why* core translation and rotation are coupled: When the core moves through the vacuum, the vacuum strands and the core effectively push against each other, due to their impenetrability. The result is a motion that resembles the motion of an asymmetrical body in a viscous fluid.

When an asymmetrical body is moved through a viscous fluid, it starts rotating. For example, this happens when a stone falls through water or honey. The rotation results
Ref. 163 from the asymmetrical shape of the body. All the tangle cores of elementary particles are asymmetrical. The strand model thus predicts that tangle cores will rotate when they move through vacuum. In other terms, the strand model predicts

▷ Linked and localized tangles have mass.
▷ Unknotted or unlinked, unlocalized tangles, such as those of photons, are predicted to be massless.

We also deduce that the more complicated a tangle is, the higher the mass value is.

In addition to the geometry effect due to the core, which is valid for massive bosons and fermions, the rotation of fermions is also influenced by the tails. The effective volume required by the belt trick will influence the coupling between translation and rotation. This effective volume will depend on the topology of the tangle core, and on the number of its tails. We again deduce that, for a given number of tails, a complicated core topology
Page 358 implies a high mass value.

In other words, the strand model links the mass m of a particle to its tangle topology: *large tangle cores have large mass.* The strand model thus predicts

▷ Particle masses are *calculable* – if the tangle topology is known.

This is an exciting prospect! To sum up, the strand model predicts that experiments in viscous fluids can lead to a deeper understanding of the masses of elementary particles.

The tangle model also implies that the mass of elementary particles – thus of particles
Challenge 137 e made of few strands – will be much smaller than the Planck mass. This is the first hint
Page 360 that the strand model solves the so-called *mass hierarchy problem* of particle physics.

At this point, however, we are still in the dark about the precise origin of particle mass values. We do not know how to calculate them. Nevertheless, the missing steps are clear: first, we need to determine the tangle topology for each elementary particle; then we need to deduce their mass values, i.e., the relation between their rotation and translation. This is a central aim in the following.

An example of the issues that arise: How does the mass value depend on the number of strands in a tangle? How does mass depend on the type of tangle?

Challenge 138 e

Potentials

In quantum mechanics, interactions are described by potentials. An *electric potential* $V(x, t)$ changes the total energy of a particle with charge q at position x, since in quantum mechanics, electric potentials influence the rotation velocity of the wave function. As a result, with an electric potential, the left-hand side of the Schrödinger equation (128), the energy term, is changed from $\hbar\omega\psi(x, t)$ to $(\hbar\omega - qV)\psi(x, t)$.

Another possibility is a potential that does not change the rotation velocity, but that changes the wavelength of a charged particle. Such a *magnetic vector potential* $A(x, t)$ thus changes the momentum term $\hbar\boldsymbol{k}$ on the right-hand side of Schrödinger's equation to $(\hbar\boldsymbol{k} - q\boldsymbol{A})\psi(x, t)$. This double substitution, the so-called *minimal coupling*, is equivalent to the statement that quantum electrodynamics has a U(1) gauge symmetry. We will deduce it in detail in the next chapter.

Page 235

In the strand model of quantum mechanics, potentials are introduced in precisely the same way as in usual quantum mechanics, so that the full Schrödinger equation for charged particles in external fields is recovered:

$$(i\hbar\partial_t - qV)\psi = \frac{1}{2m}(-i\hbar\boldsymbol{\nabla} - q\boldsymbol{A})^2\psi \,. \qquad (129)$$

This equation is the simplest formulation of quantum theory. We saw in the fourth volume that it describes and explains the size of atoms and molecules, and thus of all objects around us; and we saw that it also explains the (relative) colours of all things. The equation also explains interference, tunnelling and decay.

In summary, a non-relativistic fluctuating tangle reproduces the full Schrödinger equation. An obvious question is: how does the strand model explain the influence of interactions on the rotation speed and on the wavelength of tangles? In other words: why do strands imply minimal coupling? We will answer this question in the following chapter, on gauge interactions.

Page 225

Quantum interference from tangles

The observation of *interference* of quantum particles is due to the linear combination of states with different phases at the same position in space. Tangle functions, being wave functions, reproduce the effect. But again, it is both more fun and more instructive to explain and visualize interference with the help of tangles.

Page 191

As mentioned above, a pure change of phase of a state ψ is defined by multiplication by a complex number of unit norm, such as $e^{i\beta}$. This corresponds to a rotation of the tangle core by an angle 2β, where the factor 2 is due to the belt trick of Figure 19.

Page 177

To deduce interference, we simply use the above definition of linear combinations of tangles. This leads to the result shown in Figure 36. We find, for example, that a symmetric sum of a tangle and the same tangle with the phase rotated by $\pi/2$ (thus a core rotated by π) results in a tangle whose phase is rotated by the intermediate angle,

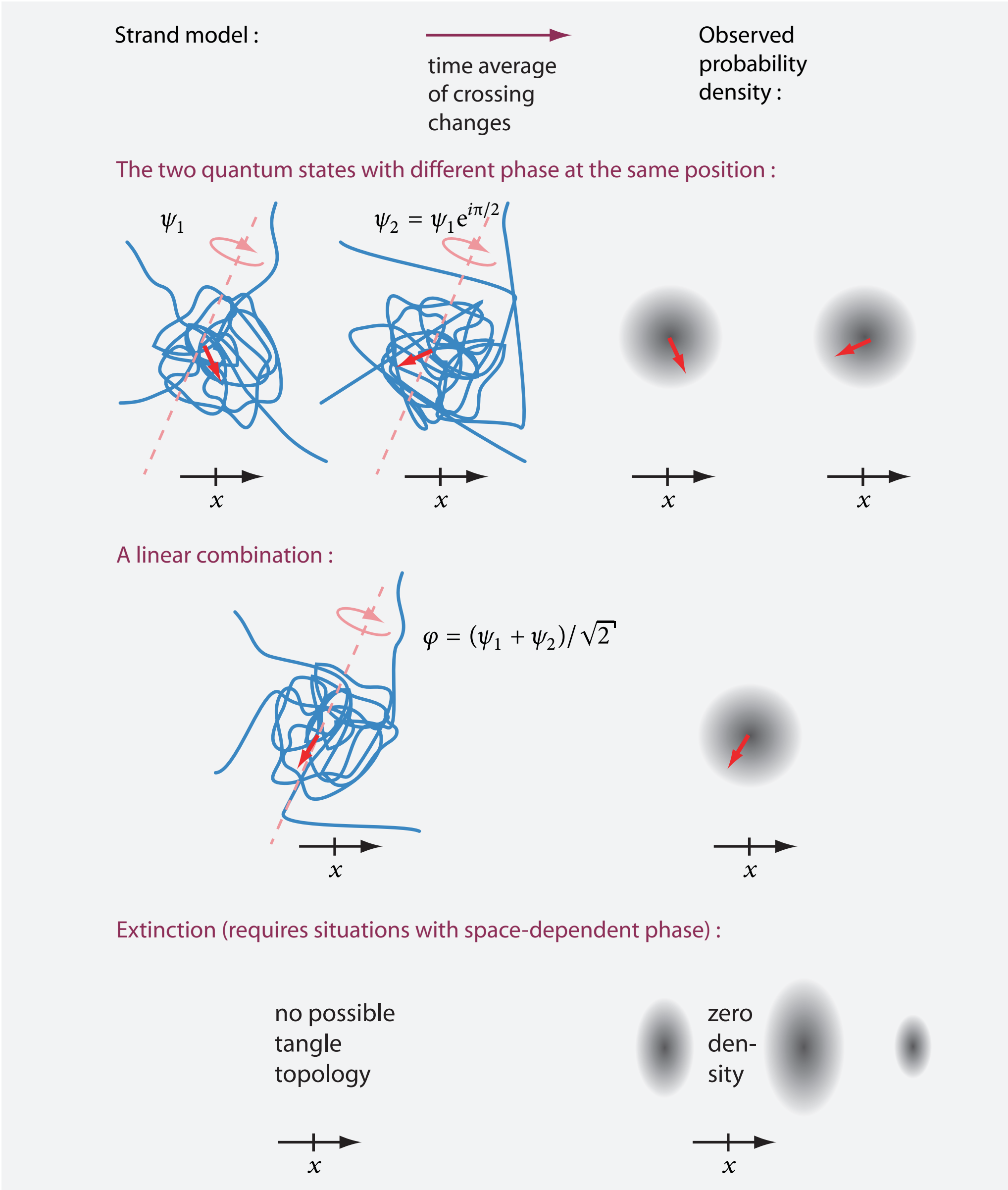


FIGURE 36 Interference: the linear combination of strands with different phase, but located at the same position.

thus $\pi/4$.

The most interesting case of interference is that of *extinction*. Scalar multiplication of a tangle function ψ by -1 gives the negative of the tangle function, the additive inverse $-\psi$. The sum of a tangle function with its negative is zero. This gives extinction in usual quantum theory. Let us check the result in the strand model, using the tangle definition of linear combinations. We have seen above that the negative of a tangle is a tangle whose core is rotated by 2π. Using the tangle definition of linear combination, we find that it is *topologically impossible* to draw or construct a localized tangle for the sum of a quantum state with its negative. The resulting particle tangle therefore must have vanishing crossing density in spatial regions where this operation is attempted. In short, particle tangles do explain extinction. And as expected from quantum particles, the explanation of extinction directly involves the tangle structure.

Deducing the Pauli equation from tangles

As we have seen, the Schrödinger equation describes the motion of quantum particles when their spin is neglected, by assuming that spin is constant over space and time. The next step is thus to include the variations of spin over space and time. This turns out to be quite straightforward.

In the strand model, spin is modelled by the continuous rotation of a tangle. We also saw that we get wave functions from tangles if we average over short time scales. At a given position in space, a tangle function will have a local average density of crossings, a local average phase, and new, a local average orientation of the rotation axis of the tangle.

To describe the axis and orientation of the tangle core, we use the *Euler angles* α, β and γ. This yields a description of the tangle function as
Ref. 164

$$\Psi(x,t) = \sqrt{\rho}\, e^{i\alpha/2} \begin{pmatrix} \cos(\beta/2)e^{i\gamma/2} \\ i\sin(\beta/2)e^{-i\gamma/2} \end{pmatrix} , \qquad (130)$$

which is the natural description of a tangle that includes the orientation of the axis. As before, the crossing density is the square root of the probability density $\rho(x,t)$. The angle $\alpha(x,t)$, as before, describes the phase, i.e., (one half of) the rotation *around* the axis. The local orientation of the axis is described by a two-component matrix and uses the two angles $\beta(x,t)$ and $\gamma(x,t)$. Due to the belt trick, the expression for the tangle function only contains *half* angles. And indeed, due to the half angles, the two-component matrix is not a vector, but a *spinor*. (The term 'spinor' was coined by well-known physicist Paul Ehrenfest in analogy to 'vector' and 'tensor'; the English pronunciation is 'spinnor'.) For $\beta = \gamma = 0$, the previous wave function ψ is recovered.

The other ingredient we need is a description of the spinning motion of the tangle. In contrast to the Schrödinger case, the spinning motion itself must be added in the description. A spinning tangle implies that the propagation of the wave is described by the wave vector $\boldsymbol{k}$ multiplied with the spin operator $\boldsymbol{\sigma}$. The *spin operator* $\boldsymbol{\sigma}$, for the wave function just given, is defined as the vector of three matrices

$$\boldsymbol{\sigma} = \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right) . \qquad (131)$$

The three matrices are the well-known *Pauli matrices.*

We now take the description of the axis orientation and the description of the spinning and insert both, as we did for the Schrödinger equation, into the non-relativistic dispersion relation $\hbar\omega = E = p^2/2m = \hbar^2k^2/2m$. We then get the wave equation

$$i\hbar\partial_t\Psi = -\frac{\hbar^2}{2m}(\boldsymbol{\sigma}\nabla)^2\Psi\ . \tag{132}$$

This is *Pauli's equation* for the evolution of a free quantum particle with spin 1/2.

As final step, we include the electric and the magnetic potentials, as we did in the case of the Schrödinger equation. We again use *minimal coupling*, substituting $i\hbar\partial_t$ by $i\hbar\partial_t - qV$ and $-i\hbar\nabla$ by $-i\hbar\nabla - q\boldsymbol{A}$, thus introducing electric charge q and the potentials

Challenge 139 s

V and $\boldsymbol{A}$. A bit of algebra involving the spin operator then leads to the famous complete form of the Pauli equation

$$(i\hbar\partial_t - qV)\Psi = \frac{1}{2m}(-i\hbar\nabla - q\boldsymbol{A})^2\Psi - \frac{q\hbar}{2m}\boldsymbol{\sigma}\boldsymbol{B}\Psi\ , \tag{133}$$

where now the magnetic field $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ appears explicitly. The equation is famous for describing, among others, the motion of silver atoms, which have spin 1/2, in the Stern–

Vol. IV, page 83

Gerlach experiment. This is due to the new, last term on the right-hand side, which does not appear in the Schrödinger equation. The new term is a pure spin effect and predicts a g-factor of 2. Depending on the spin orientation, the sign of the last term is either positive or negative; the term thus acts as a spin-dependent potential. The two options for the spin orientation then produce the upper and the lower beams of silver atoms that are observed in the Stern–Gerlach experiment.

In summary, a non-relativistic tangle that rotates continuously reproduces the Pauli equation. In particular, such a tangle predicts that the g-factor of an elementary charged fermion is 2.

Rotating arrows and path integrals

Another simple way to visualize the equivalence between the strand model and the Pauli equation uses the formulation of quantum theory with path integrals. We recall that tangle tails are not observable, and that the tangle core defines the position and phase of the quantum particle. If the core is approximated to be of vanishing size, thus 'point-like', then the motion of the core describes the 'path' of the particle. (This equivalence was already mentioned188 above.)

Feynman described the motion of quantum particles in his famous popular book on

Ref. 165

QED as advancing rotating arrows. The continuous rotation of the tangle core visualizes Feynman's rotating little arrow. The different possible motions of the 'point-like' tangle core corresponds to different paths. Quantum theory appears when the effects of all possible paths are superposed. In particular, the phase and amplitude for each path must added like small vectors. In the strand model the effects of all possible paths are added automatically, through the fluctuations of the tangles motion. And through the definitions given above, the addition occurs in exactly the way that Feynman described.

In this way, the tangle model reproduces the path integral formulation of quantum mechanics.

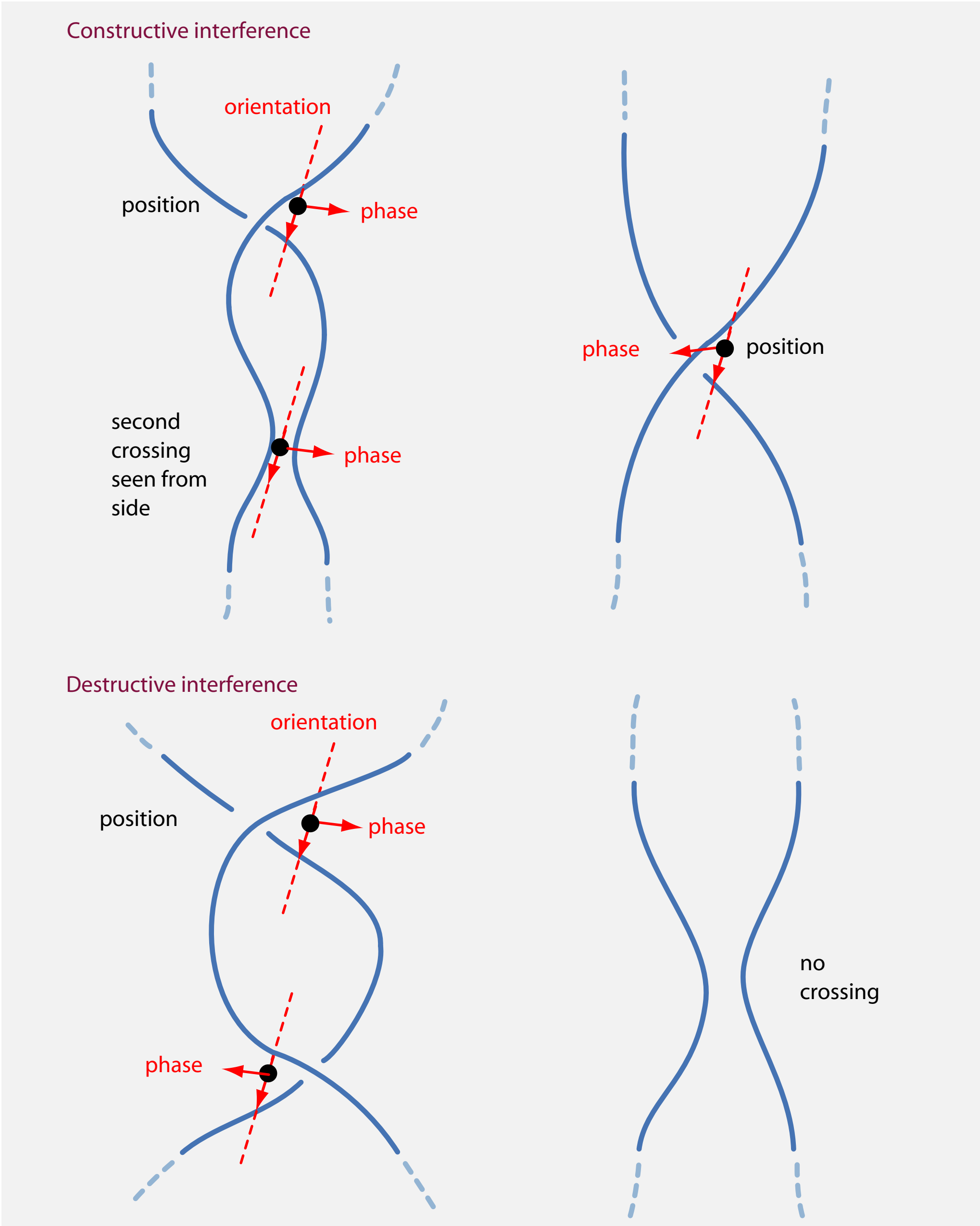


FIGURE 37 Interference in the strand model: two connected crossings – in this case with the same distance, i.e., with the same 'amplitude' – superpose constructively (top) and destructively (bottom).

Interference and double slits

Often, interference is seen as the essence of quantum theory, or as the biggest difference between quantum theory and classical physics. Also interference can be visualized with strands.

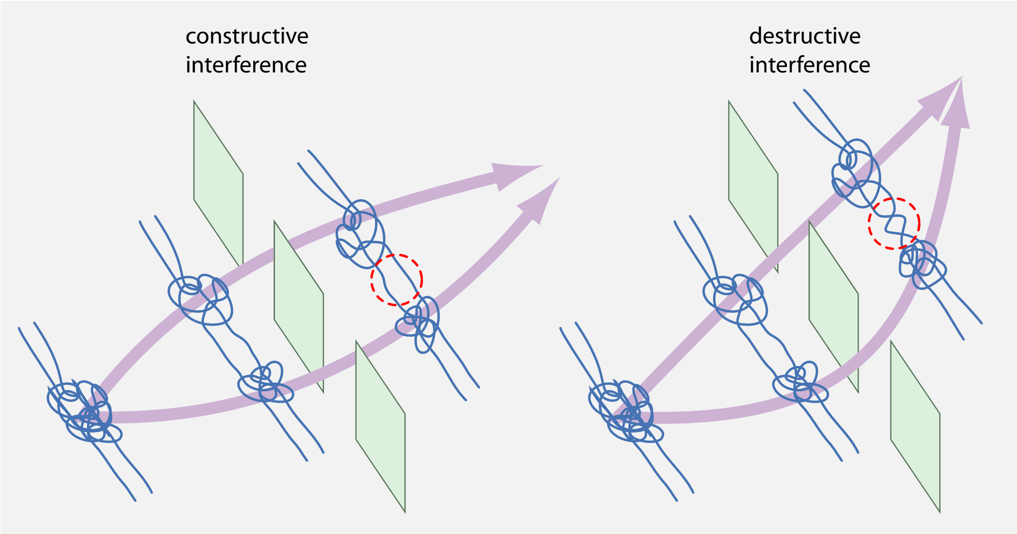


FIGURE 38 A fermion tangle passing a double slit: constructive interference (left) and destructive interference (right).

In nature, interference – whether constructive or destructive or partial – results from the superposition of states with different phase values at a point in space. For strands, superposition occurs for strands, when strand segments are connected. The phase value is defined by the crossing geometry, as explained in Figure 31. In the strand model, 'at a point in space' becomes 'at two points within Planck scale distance'. Together, this yields the fundamental superposition mechanism shown in Figure 37. The phases of two crossings on two connected strands can add up or cancel. The general case, with different amplitudes, is easily deduced.

Page 187

Challenge 140 e

The strand explanation of interference allows to describe the double slit experiment. Because of its tails, a fermion tangle obeys spinor statistics and spinor rotation behaviour. This leads to the observed interference behaviour for spin 1/2 particles, as visualized in Figure 38. The corresponding visualization for photon interference is given in Figure 39.

Interference is essential in all effects of quantum physics. We explore a few additional ones.

Measurements and wave function collapse

In nature, a measurement of a quantum system in a superposition is observed to yield one of the possible eigenvalues and to prepare the system in the corresponding eigenstate. In nature, the probability of each measurement outcome depends on the coefficient of that eigenstate in the superposition.

Vol. IV, page 143

To put the issue into context, here is a short reminder from quantum mechanics. Every measurement apparatus shows measurement results. Thus, every measurement apparatus is a device with memory. (In short, it is *classical*.) All devices with memory contain one or several baths. Thus, every measurement apparatus couples at least one bath to the system it measures. The coupling depends on and defines the observable to be measured by the apparatus. Every coupling of a bath to a *quantum* systems leads to decoherence.

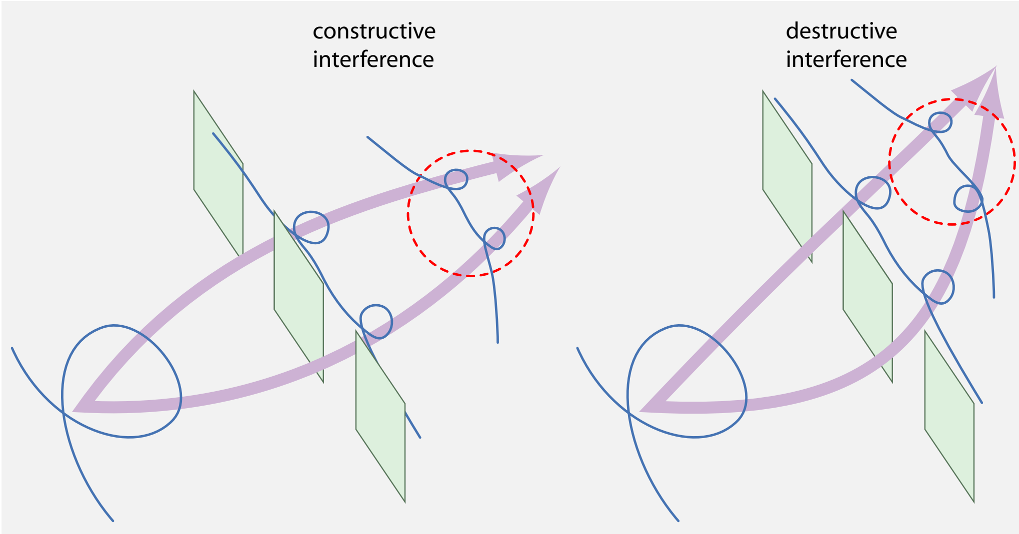


FIGURE 39 The double-slit experiment with photons: constructive interference (left) and destructive interference (right).

Decoherence leads to probabilities and wave function collapse. In short, collapse and measurement probabilities are necessary and automatic in quantum theory.

The strand model describes the measurement process in precisely the same way as standard quantum theory; in addition, it *visualizes* the process.

▷ A ***measurement*** is modelled as a strand deformation induced by the measurement apparatus that 'pulls' a tangle towards the resulting eigenstate.

▷ This pulling of strands models and visualizes the ***collapse*** of the wave function.

An example of measurement is illustrated in Figure 40. When a measurement is performed on a superposition, *the untangled 'addition region' can be imagined to shrink into disappearance.* For this to happen, one of the underlying eigenstates has to 'eat up' the other: that is the collapse of the wave function. In the example of the figure, the addition region can disappear either towards the outside or towards the inside. The choice is due to the bath that is coupled to the system during measurement; the bath thus determines the outcome of the measurement. We also deduce that the probability of measuring a particular eigenstate will depend on the (weighed) volume that the eigenstate took up in the superposition.

This visualization of the wave function collapse also makes clear that the collapse is not limited by any speed limit, as no energy and no information is transported. Indeed, the collapse happens by displacing strands and at most crossings, but does not produce any crossing changes.

In summary, the strand model describes measurements in precisely the same way as usual quantum theory. In addition, *strands visualize the collapse of the wave function as a shape deformation from a superposed tangle to an eigenstate tangle.*

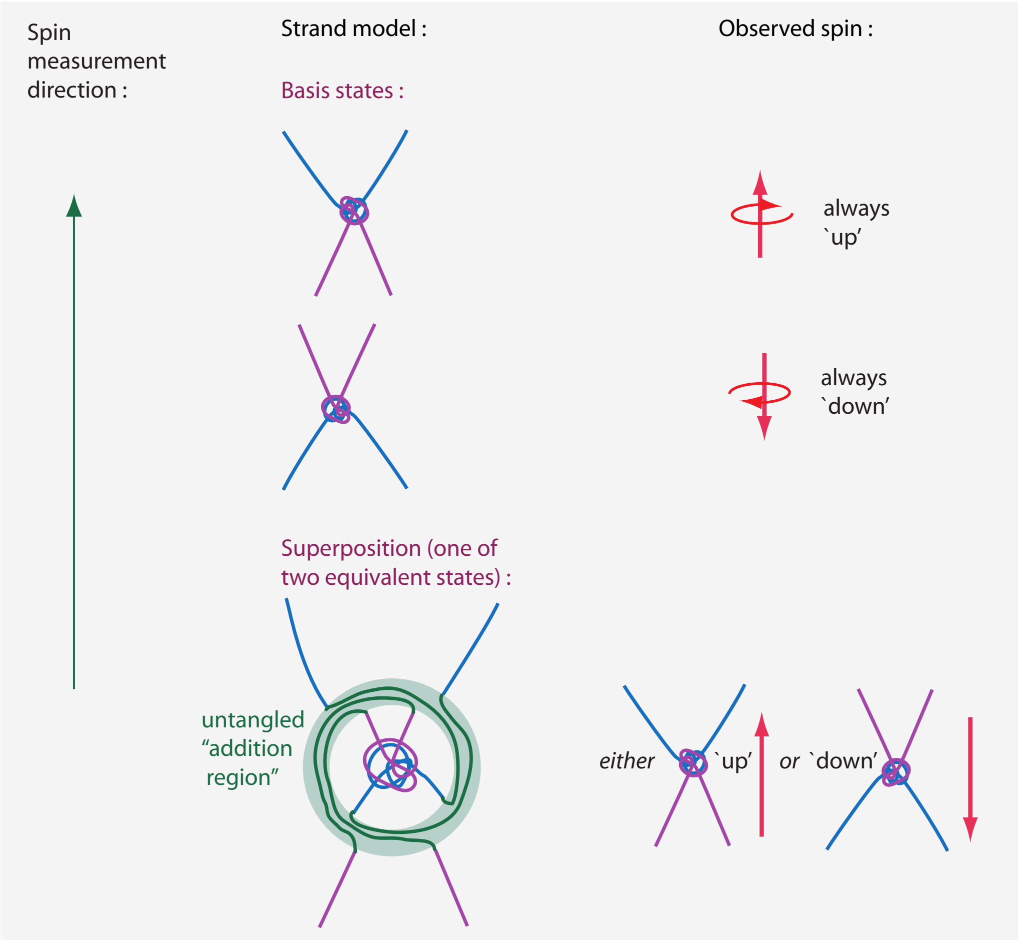


FIGURE 40 Measurement of a spin superposition: the addition region disappears either outwards or inwards.

Hidden variables and the Kochen–Specker theorem

At first sight, the strand model seems to fall into the trap of introducing hidden variables into quantum theory. One could indeed argue that the shapes (and fluctuations) of the strands play the role of hidden variables. On the other hand, it is well known that non-contextual hidden variables are impossible in quantum theory, as shown by the Kochen–Specker theorem (for sufficiently high Hilbert-space dimensions). Is the strand model flawed? No.
Ref. 166

We recall that strands are not observable. In particular, strand shapes are not physical observables and thus not physical (hidden) variables either. Even if we tried promoting strand shapes to physical variables, the evolution of the strand shapes would only be observable through the ensuing crossing switches. And crossing switches evolve due to the influence of the environment, which consists of all other strands in nature, including those of space-time itself. Thus

▷ The evolution of strand shapes and crossing switches is *contextual*.

Therefore, the strand model does not contradict the Kochen–Specker theorem.

In simple language, in quantum theory, hidden variables are not a problem if they are properties of the environment, and not of the quantum system itself. This is precisely the case for the strand model. For a quantum system, the strand model provides no hidden variables. In fact, for a quantum system, the strand model provides no variables beyond the usual ones from quantum theory. And as expected and required from any model that reproduces decoherence, the strand model leads to a contextual, probabilistic description of nature.

In summary, despite using fluctuating tangles as underlying structure, the strand model is equivalent to usual quantum theory. The strand model contains nothing more and nothing less than usual quantum theory.

Many-particle states and entanglement

In nature, the quantum states of two or more particles can be *entangled*. Entangled states are many-particle states that are not separable. Entangled states are one of the most fascinating quantum phenomena; especially in the case of macroscopic entanglement, they are still being explored in many experiments. We will discover that the strand model visualizes them simply and clearly.

To describe entanglement, we first need to clarify the notion of many-particle state. In the strand model,

▷ A ***many-particle state*** is composed of several tangles.

In this way, an N-particle wave function defines N values at every point in space, one value for each particle. This is possible, because in the strand model, the strands of one particle tangle are *separate* from the strands of other particles.

Usually, a N-particle wave function is described by a single-valued function in $3N$ dimensions. It is less known that a single-valued N-particle wave function in $3N$ dimensions is mathematically equivalent to an N-valued wave function in three dimensions. Usually, N-valued functions are not discussed; we feel uneasy with the concept. But the strand model naturally defines N wave function values at each point in space: each particle has its own tangle, and each tangle yields, via short-term averaging, one complex value, with magnitude and phase, at each point in space. In this way, the strand model is able to describe N particles in just 3 dimensions.

In other words, the strand model does not describe N particles with 1 function in $3N$ dimensions; it describes many-particle states with N functions in 3 dimensions. In this way, the strand model remains as close to everyday life as possible. Many incorrect statements on this issue are found in the research literature; many authors incorrectly claim the impossibility of many-particle quantum theory in 3 dimensions. Some authors even claim, in contrast to experiment, that it is impossible to visualize many-particle states in 3 dimensions. These arguments all fail to consider the possibility to define completely separate wave functions for each particle in three dimensions. (It must be said that this unusual possibility is hard to imagine if wave functions are described as

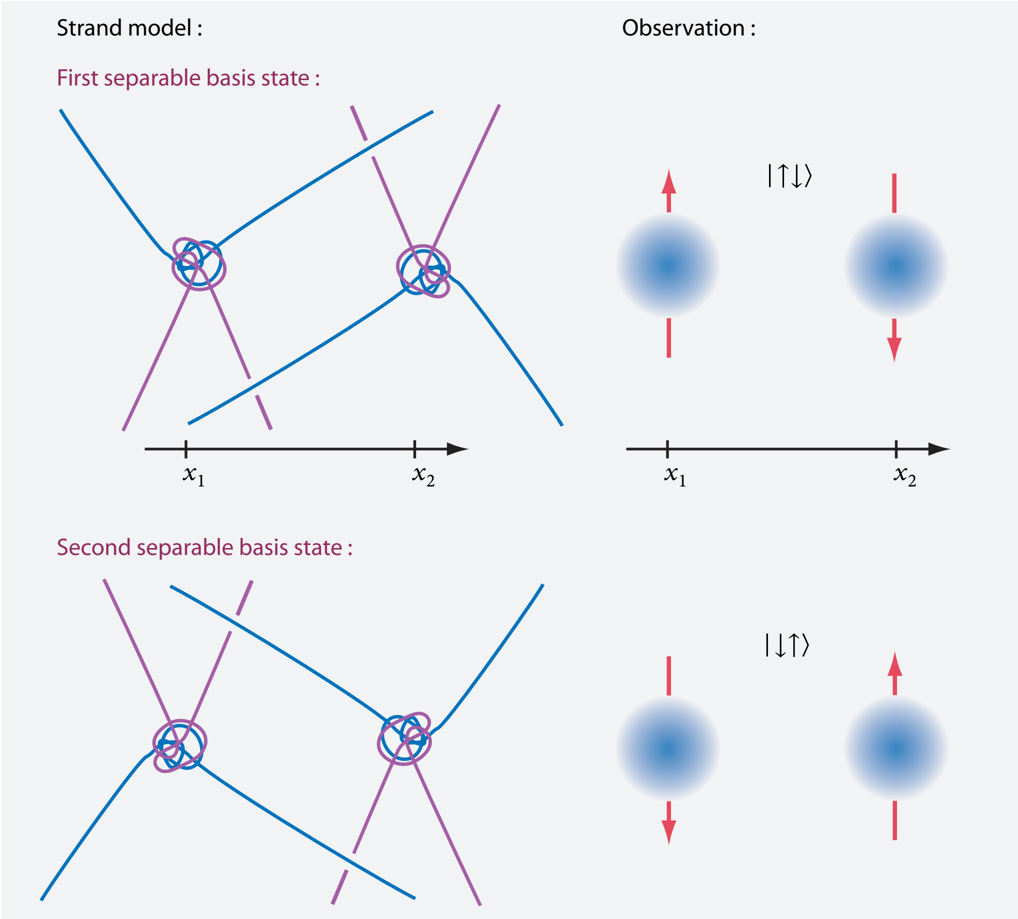


FIGURE 41 Two examples of two distant particles with spin in separable states: observation and strand model.

Ref. 165 continuous functions.) However, clear thinkers like Richard Feynman always pictured many-particle wave functions in 3 dimensions. Also in this domain, the strand model provides an underlying picture to Feynman's approach. This is another situation where the strand model eliminates incorrect thinking habits and supports the 'naive' view of quantum theory.

Now that we have defined many-particle states, we can also define entangled states.

▷ An ***entangled state*** is a non-separable superposition of separable many-particle states. State are separable when their tangles can be pulled away without their tails being entangled.

We will now show that the above definitions of superpositions and of measurements using strands are sufficient to describe entanglement.

As first example, we explore entangled states of the spin of two distant massive fermions. This is the famous thought experiment proposed by David Bohm. In the strand

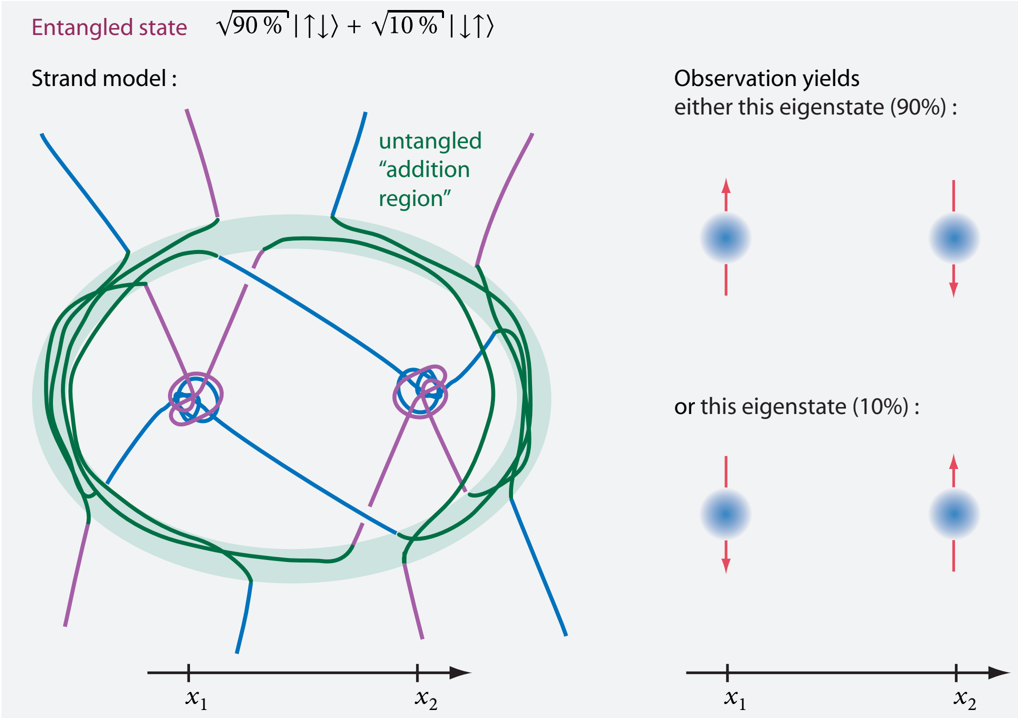


FIGURE 42 An entangled spin state of two distant particles.

model, two distant particles with spin 1/2 in a *separable* state are modelled as two distant, separate tangles of identical topology. Figure 41 shows two separable basis states, namely the two states with total spin 0 given by $|\uparrow\downarrow\rangle$ and by $|\downarrow\uparrow\rangle$. Such states can also be produced in experiments. We note that to ensure total spin 0, the tails must be imagined to cross somewhere, as shown in the figure.

We can now draw a superposition $\sqrt{90\,\%}\,|\uparrow\downarrow\rangle + \sqrt{10\,\%}\,|\downarrow\uparrow\rangle$ of the two spin-0 basis states. We simply use the definition of addition and find the state shown in Figure 42. We can now use the definition of measurement to check that the state is indeed entangled. If we measure the spin orientation of one of the particles, the untangled addition region disappears. The result of the measurement will be either the state on the inside of the addition region or the state on the outside. And since the tails of the two particles are linked, after the measurement, independently of the outcome, the spin of the two particles will always point in opposite directions. This happens for every particle distance. Despite this extremely rapid and apparently superluminal collapse, no energy travels faster than light. The strand model thus reproduces exactly the observed behaviour of entangled spin 1/2 states.

A second example is the entanglement of two photons, the well-known Aspect experiment. Also in this case, entangled spin 0 states, i.e., entangled states of photons of opposite helicity (spin), are most interesting. Again, the strand model helps to visualize the situation. Here we use the strand model for the photon that we will deduce only later on. Figure 43 shows the strand model of the two separable basis states and the strand model of the entangled state. Again, the measurement of the helicity of one photon in Page 230

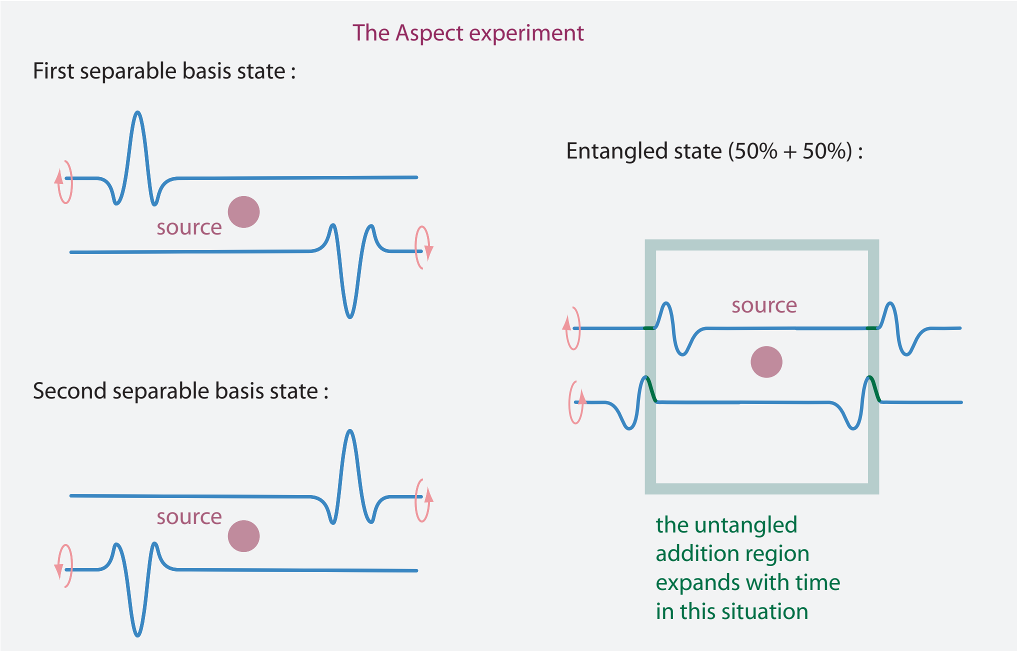


FIGURE 43 The basis states and an entangled state of two distant photons travelling in opposite directions, with total spin 0.

the entangled state will lead to one of the two basis states. And as soon as the helicity of one photon is measured, the helicity of its companion collapses to the opposite value, whatever the distance! Experimentally, the effect has been observed for distances of many kilometres. Again, despite the extremely rapid collapse, no energy travels faster than light. And again, the strand model completely reproduces the observations.

Ref. 167

Mixed states

Mixed states are statistical ensembles of pure states. In the strand model,

▷ A ***mixed state*** is a (weighted) temporal alternation of pure states.

Mixed states are important in discussions of thermodynamic quantities. We mention them to complete the equivalence of the states that appear in quantum theory with those provided by the strand model. We do not pursue this topic any further.

The dimensionality of space-time

'Nature consists of particles moving in empty space.' Democritus stated this 2500 years ago. Today, we know that is a simplified description of one half of physics: it is a simplified description of quantum theory. In fact, Democritus' statement, together with strands, allows us to argue that physical space must have three dimensions, as we will see now.

Deducing the dimensionality of physical space from first principles is an old and difficult problem. The difficulty is also due to the lack of alternative descriptions of nature. Page 169 Our exploration of the foundations of the strand model has shown that humans, animals and machines always use three spatial dimensions to describe their environment. They cannot do otherwise. Humans, animals and machines cannot talk and think without three dimensions as background space.

But how can we show that *physical space* – not the *background space* we need for thinking – is three-dimensional and must be so? We need to show that (1) all experiments reproduce the result and that (2) no other number of dimensions yields a consistent description of nature.

In nature, and also in the strand model, as long as particles can be defined, they can be rotated around each other and they can be exchanged. No experiment has ever been performed or has ever been proposed that changes this observation. The observed properties of rotations, of spin 1/2, of particle exchange and all other observations confirm that space has three dimensions. Fermions only exist in three dimensions. In the strand model, the position and the orientation of a particle is intrinsically a three-dimensional quantity; physical space is thus three-dimensional, in all situations where it can be defined. (The only situations where this definition is impossible are horizons and the Planck scales.) In short, both nature and the strand model are found to be three-dimensional at all experimentally accessible energy scales. Conversely, detecting an additional spatial dimension would directly invalidate the strand model.

Nature has three dimensions. The only way to predict this result is to show that no other number is possible. The number of dimensions of nature can only result from a self-consistency argument. And interestingly, the strand model produces such an argument.

In the strand model, knots and tangles are impossible to construct in physical spaces with dimensions *other* than three. Indeed, mathematicians can show that in four spatial dimensions, every knot and every tangle can be undone. (In this argument, time is not and does not count as a fourth spatial dimension, and strands are assumed to remain one-dimensional entities.) Worse, in the strand model, spin does not exist in spaces that have more or fewer than three dimensions. Also the vacuum and its quantum fluctuations do not exist in more than three dimensions. Moreover, in other dimensions it is impossible to formulate the fundamental principle. In short, the strand model of matter and of observers, be they animals, people or machines, is possible in three spatial dimensions only. No description of nature with a background or physical space of more or less than three dimensions is possible with strands. Conversely, constructing such a description would invalidate the strand model.

The same type of arguments can be collected for the one-dimensionality of physical Challenge 141 e time. It can be fun exploring them – for a short while. In summary, the strand model *only* works in $3+1$ space-time dimensions; it does not allow any other number of dimensions. Page 164 We have thus ticked off another of the millennium issues. We can thus continue with our adventure.

Operators and the Heisenberg picture

In quantum theory, Hermitean operators play an important role. In the strand model, *Hermitean* or self-adjoint operators are operators that leave the tangle topology invariant.

Also unitary operators play an important role in quantum theory. In the strand model, *unitary* operators are operators that deform tangles in a way that the corresponding wave function retains its norm, i.e., such that tangles retain their topology and their core shape.

Physicists know two ways to describe quantum theory. One is to describe evolution with time-dependent quantum states – the *Schrödinger picture* we are using here – and the other is to describe evolution with time-dependent operators. In this so-called *Heisenberg picture*, the temporal evolution is described by the operators.

The two pictures of quantum theory are equivalent. In the Heisenberg picture, the fundamental principle, the equivalence of a crossing switch with $\hbar$, becomes a statement
Ref. 168 on the behaviour of operators. Already in 1987, Louis Kauffman had argued that the commutation relation for the momentum and position operators

$$px - xp = \hbar i \tag{134}$$

is related to a crossing switch. The present section confirms that speculation.

In quantum mechanics, the commutation relation follows from the definition of the momentum operator as $p = \hbar k$, $k = -i\partial_x$ being the wave vector operator. The factor $\hbar$ defines the unit of momentum. The wave vector counts the number of wave crests of a wave. Now, in the strand model, a rotation of a state by an angle π is described by a multiplication by i. Counting wave crests of a propagating state is only possible by using the factor i, as this factor is the only property that distinguishes a crest from a trough. In short, the commutation relation follows from the fundamental principle of the strand model.

Lagrangians and the principle of least action

Before we derive the Dirac equation, we show that the strand model naturally leads to describe motion with Lagrangians.

In nature, physical action is an observable measured in multiples of the natural unit, the quantum of action $\hbar$. Action is the fundamental observable about nature, because *action measures the total change occurring in a process.*

In the strand model,

▷ The physical ***action*** W of a physical process is the observed number of crossing switches of strands. Action values are multiples of $\hbar$.

We note that these multiples, if averaged, do not need to be integer multiples. We further note that through this definition, *action is observer-invariant.* This important property is thus automatic in the strand model.

In nature, energy is action per time. Thus, in the strand model we have:

▷ ***Energy*** is the number of crossing switches per time in a system.

In nature, when free quantum particles move, their phase changes linearly with time. In other words, the 'little arrow' representing the free particle phase rotates with constant angular frequency. We saw that in the strand model, the 'little arrow' is taken as (half) the orientation angle of the tangle core, and the arrow rotation is (half) the rotation of the tangle core.

▷ The ***kinetic energy*** T of a particle is the number of crossing switches per time induced by shape fluctuations of the continuously rotating tangle core.

We call $\mathcal{T}$ the corresponding volume density: $\mathcal{T} = T/V$. In nature, the Lagrangian is a practical quantity to describe motion. For a *free* particle, the Lagrangian density $\mathcal{L} = \mathcal{T}$ is simply the kinetic energy density, and the action $W = \int \mathcal{L}\ \mathrm{d}V\mathrm{d}t = Tt$ is the product of kinetic energy and time. In the strand model, a free particle is a constantly rotating and advancing tangle. We see directly that this constant evolution minimizes the action W for a particle, given the states at the start and at the end.

This aspect is more interesting for particles that interact. Interactions can be described by a potential energy U, which is, more properly speaking, the energy of the field that produces the interaction. In the strand model,

▷ ***Potential energy*** U is the number of crossing switches per time induced by an interaction field.

We call $\mathcal{U}$ the corresponding volume density: $\mathcal{U} = U/V$. In short, in the strand model, an interaction changes the rotation rate and the linear motion of a particle tangle.

In the strand model, the *difference* between kinetic and potential energy is thus a quantity that describes how much a system consisting of a tangle and a field *changes* at a given time. The total change is the integral over time of all instantaneous changes. In other words, in the strand model we have:

▷ The ***Lagrangian density*** $\mathcal{L} = \mathcal{T} - \mathcal{U}$ is the number of crossing switches per volume and time, averaged over many Planck scales.

▷ The physical ***action*** $W = \int L\,\mathrm{d}t = \int\int \mathcal{L}\ \mathrm{d}V\mathrm{d}t$ of a physical process is the observed number of crossing switches of strands. The action value W_{if} between an initial state ψ_{i} and a final state ψ_{f} is given by

$$W_{\mathrm{if}} = \langle \psi_{\mathrm{i}} \,|\int \mathcal{L}\,\mathrm{d}t \,|\, \psi_{\mathrm{f}} \rangle = \langle \psi_{\mathrm{i}} \,|\int (\mathcal{T} - \mathcal{U})\,\mathrm{d}t \,|\, \psi_{\mathrm{f}} \rangle \ . \tag{135}$$

Since energy is related to crossing switches, it is natural that strand fluctuations that do *not* induce crossing switches are *favoured*. In short, the strand model states

▷ Evolution of strands *minimizes crossing switch number.* As a result, strands minimize the action W.

In the strand model, *the least action principle appears naturally*. In the strand model, an evolution has least action when it occurs with the smallest number of crossing changes. With this connection, one can also show that the strand model implies Schwinger's quantum action principle.

Challenge 142 e

To calculate quantum motion with the principle of least action, we need to define the kinetic and the potential energy in terms of strands. There are various possibilities for Lagrangian densities for a given evolution equation; however, all are equivalent. In case of the free Schrödinger equation, one possibility is:

$$\mathcal{L} = \frac{i\hbar}{2}(\overline{\psi}\,\partial_t\psi - \partial_t\overline{\psi}\,\psi) - \frac{\hbar^2}{2m}\nabla\overline{\psi}\nabla\psi \;. \qquad (136)$$

In this way, the principle of least action can be used to describe the evolution of the Schrödinger equation. The same is possible for situations with potentials, for the Pauli equation, and for all other evolution equations of quantum particles.

We thus retain that the strand model explains the least action principle. It explains it because strand evolution minimizes the number of crossing switches.

Special relativity: the vacuum

In nature, there is an invariant limit energy speed c, namely the speed of light and of all other massless radiation. Special relativity is the description of the consequences from this observation, in the case of a flat space-time.

We remark that special relativity also implies and requires that the flat vacuum looks exactly the same for all inertial observers. In the strand model, the idea of flat vacuum as a set of fluctuating featureless strands that are *unknotted* and *unlinked* automatically implies that for any inertial observer the flat vacuum has no matter content, has no energy content, is isotropic and is homogeneous. The strand model thus realizes this basic requirement of special relativity. In the strand model, vacuum is *Lorentz-invariant*.

Many models of the vacuum, even fluctuating ones, have difficulties reproducing Lorentz invariance. The strand model differs, because the strands are not the observable entities; only their crossing switches are. This topological definition, together with the averaging of the fluctuations, makes the vacuum Lorentz-invariant.

We note that in the strand model, the vacuum is unique, and the vacuum energy of flat infinite vacuum is exactly zero. In the strand model, there is no divergence of the vacuum energy, and there is thus *no* contribution to the cosmological constant from quantum field theory. In particular, in the strand model there is no need for supersymmetry to explain the small energy density of the vacuum.

Special relativity: the invariant limit speed

In the strand model, massless particles are unknotted and untangled. Even though we will deduce the strand model for photons only later on, we use it here already, to speed up the discussion. In the strand model, the *photon* is described by a single, helically deformed unknotted strand, as shown in Figure 51. Therefore, we can define:

Page 230

▷ The ***Planck speed*** c is the observed average speed of crossing switches due to photons.

Because the definition uses crossing switches and a massless particle, the speed of light c is an *energy speed*. Also speed of light c is an average for long times. Indeed, as is well-

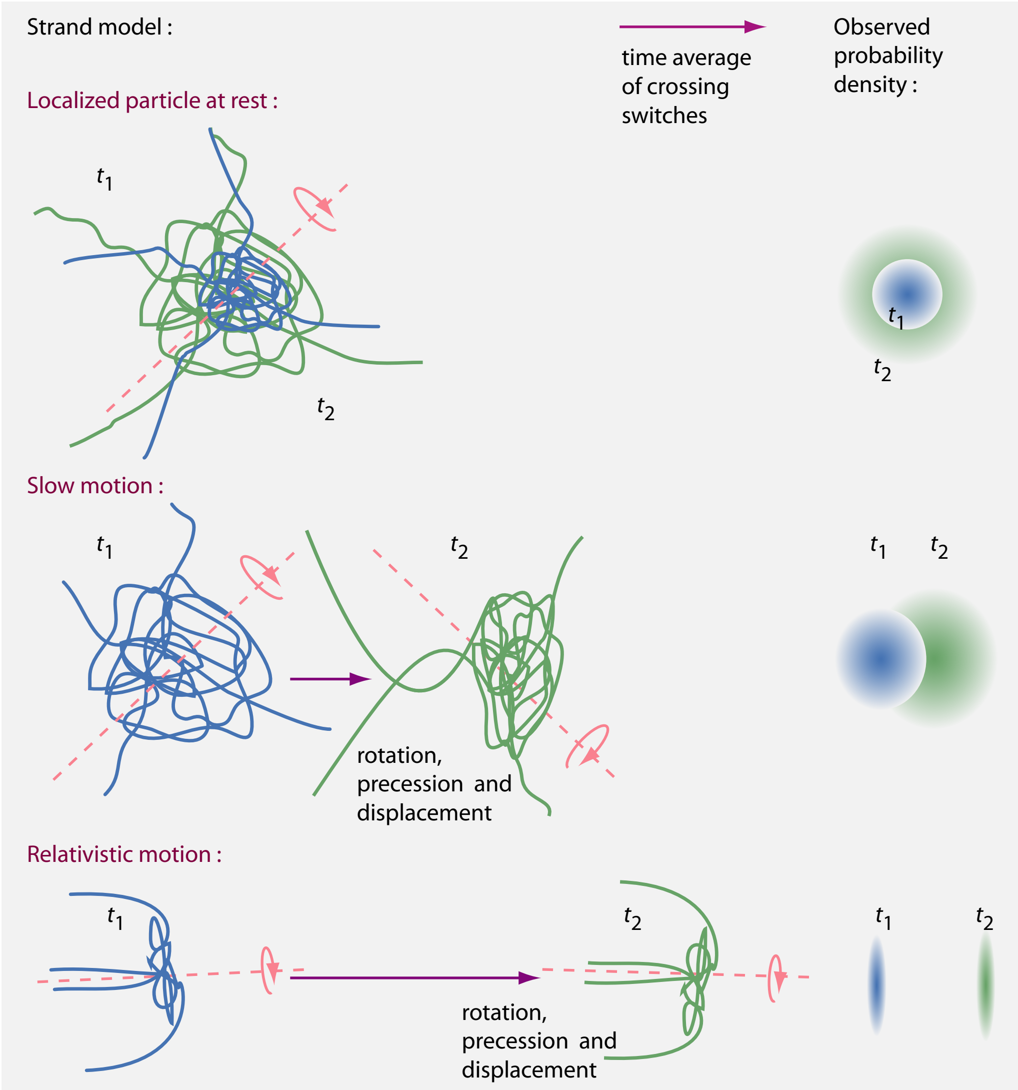


FIGURE 44 Tangles at rest, at low speed and at relativistic speed.

Ref. 169 known in quantum field theory, due to the indeterminacy relation, single photons can travel faster or slower than light, but the probability for large deviations is extremely low.

The linear motion of a helically deformed photon strand through the vacuum strands is similar to the motion of a bottle opener through cork. It differs from the linear motion of a matter tangle through vacuum, which makes use of the belt trick. The belt trick slows fermions down, though the details are not simple, as we will discover below. In short, we find that matter tangles always move *more slowly than light*. The speed c is a *limit speed*. Page 351

In fact, we see that ultrarelativistic tangles move, as shown in Figure 44, almost like light. We thus find that matter can *almost* reach the speed of light. The speed c is indeed a *limit speed* for matter.

However, one problem remains open: how exactly do tangles move through the web

Page 351 that describes the vacuum? We will clarify this issue later on. In a few words, the motion of a photon requires that the strands of the surrounding space make room for it. This requires favourable fluctuations, thus a finite time. The motion process of photons thus makes it clear that the speed of light is *finite*.

The speed of light c is defined as an average, because, as well-known in quantum field theory, there are small probabilities that light moves faster or slower that c. But the average result c will be the same for every observer. The value of the speed c is thus *invariant*.

In 1905, Einstein showed that the mentioned properties of the speed of light – energy speed, limit speed, finite speed and invariant speed – imply the Lorentz transformations. In particular, the three properties of the speed of light c imply that the energy E of a particle of mass m is related to its momentum p as

$$E^2 = m^2c^4 + c^2p^2 \quad \text{or} \quad \hbar^2\omega^2 = m^2c^4 + c^2\hbar^2k^2 \, . \tag{137}$$

This dispersion relation is thus also valid for massive particles made of tangled strands – Page 358 even though we cannot yet calculate tangle masses. (We will do this later on.)

Should we be surprised at this result? No. In the fundamental principle, the definition Page 152 of the crossing switch, we inserted the speed of light as the ratio between the Planck length and the Planck time. Therefore, by defining the crossing switch in the way we did, we have implicitly stated the invariance of the speed of light.

Fluctuating strands imply that flat vacuum has no matter or energy content, for *every* inertial observer. Due to the strand fluctuations, flat vacuum is also homogeneous and isotropic for every inertial observer. Therefore, together with the $3+1$-dimensionality of Page 209 space-time deduced above, we have now definitely shown that flat vacuum has Poincaré Page 164 symmetry. This settles another issue from the millennium list.

Page 196 The relativistic dispersion relation differs from the non-relativistic case in two ways. First, the energy scale is shifted, and now includes the rest energy $E_0 = c^2m$. Secondly, the spin precession is not independent of the particle speed any more; for relativistic particles, the spin lies close to the direction of motion. Both effects follow from the existence of a limit speed.

If we neglect spin, we can use the relativistic dispersion relation to deduce directly the well-known Klein–Gordon equation for the evolution of a wave function:

$$-\hbar^2\partial_{tt}\psi = m^2c^4 - c^2\hbar^2\nabla^2\psi \, . \tag{138}$$

In other words, the strand model implies that relativistic tangles follow the Klein–Gordon equation. We now build on this result to deduce Dirac's equation for relativistic quantum motion.

Dirac's equation deduced from tangles

The relativistic Klein–Gordon equation assumes that spin effects are negligible. This approximation fails to describe most experiments. A precise description of relativistic elementary particles must include spin.

So far, we deduced the Schrödinger equation using the relation between phase and the quantum of action, using the non-relativistic energy–momentum relation, and neglecting spin. In the next step we deduced the Pauli equation by including the properties of spin 1/2. The following step was to deduce the Klein–Gordon equation using again the relation between phase and the quantum of action, this time the relativistic energy–momentum relation, but assuming zero spin. The final and correct description of elementary fermions, the Dirac equation, results from combining all three ingredients: (1) the relation between the quantum of action and the phase of the wave function, (2) the relativistic mass–energy relation, and (3) the effects of spin 1/2. Now we can reproduce this derivation because all three ingredients are reproduced by the strand model.

We first recall the derivation of the Dirac equation found in textbooks. The main observation about spin in the relativistic context is the existence of states of right-handed and of left-handed chirality: spin can precess in two opposite senses around the direction of momentum. In addition, for massive particles, the two chiral states mix. The existence of two chiralities requires a description of spinning particles with a wave function that has *four* complex components, thus *twice* the number of components that appear in the Pauli equation. Indeed, the Pauli equation implicitly assumes only one, given sign for the chirality, even though it does not specify it. This simple description is possible because in non-relativistic situations, states of different chirality do not mix.

Consistency requires that each of the four components of the wave function of a relativistic spinning particle must follow the relativistic energy–momentum relation, and thus the Klein–Gordon equation. This requirement is known to be sufficient to deduce the Dirac equation. One of the simplest derivations is due to Lerner; we summarize it here.

Ref. 170
Ref. 171

When a spinning object moves relativistically, we must take both chiralities into account. We call u the negative chiral state and v the positive chiral state. Each state is described by two complex numbers that depend on space and time. The 4-vector for probability and current becomes

$$J_\mu = u^\dagger \sigma_\mu u + v^\dagger \sigma_\mu v \,. \tag{139}$$

We now introduce the four-component spinor φ and the 4×4 spin matrices a_μ

$$\varphi = \begin{pmatrix} u \\ v \end{pmatrix} \quad \text{and} \quad \alpha_\mu = \begin{pmatrix} \sigma_\mu & 0 \\ 0 & \overline{\sigma}_\mu \end{pmatrix} , \tag{140}$$

where $\sigma_\mu = (I, \boldsymbol{\sigma})$ and $\overline{\sigma}_\mu = (I, -\boldsymbol{\sigma})$ and I is the 2×2 identity matrix. The 4-current can then be written as

$$J_\mu = \varphi^\dagger \alpha_\mu \varphi \,. \tag{141}$$

The three requirements of current conservation, Lorentz invariance and linearity then yield the evolution equation

Ref. 171

$$\mathrm{i}\hbar \partial^\mu (\alpha_\mu \varphi) + mc\gamma_5 \varphi = 0 \,. \tag{142}$$

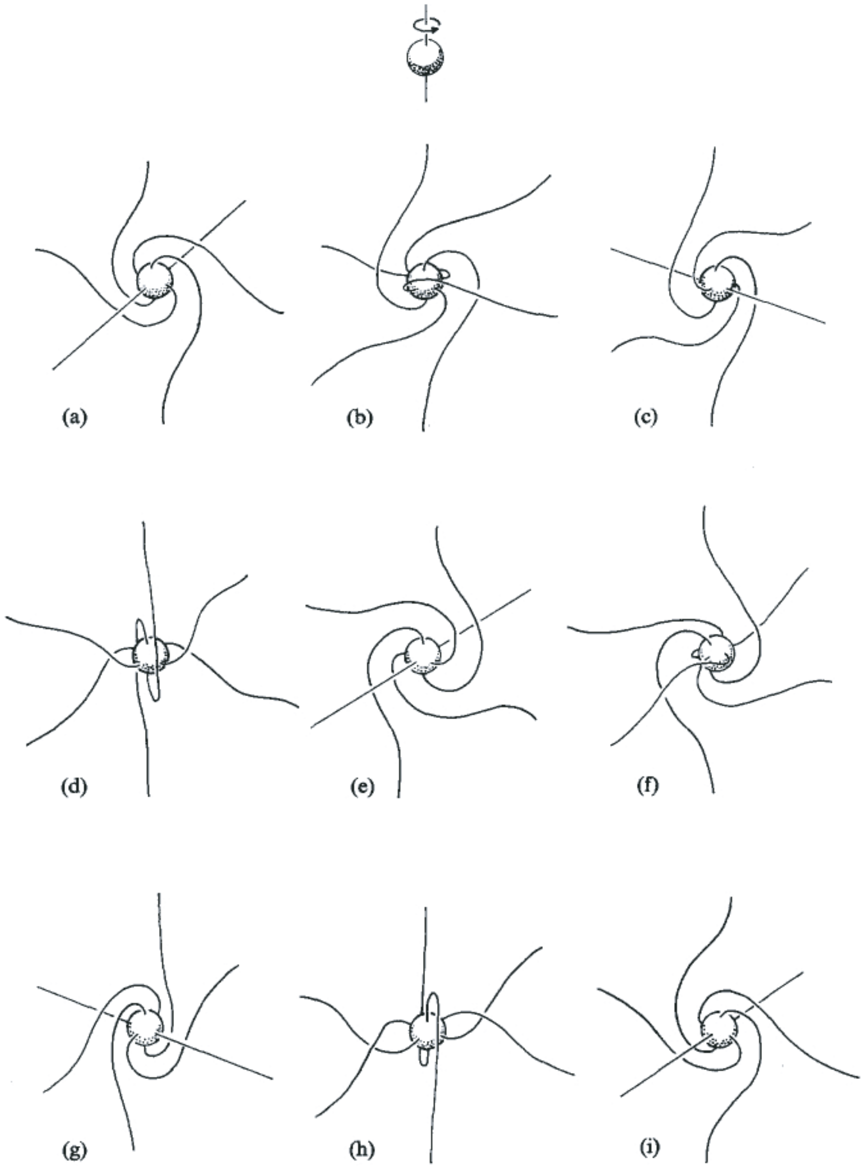


FIGURE 45 The belt trick for a rotating body with many tails, as used by Battey-Pratt and Racey to deduce the Dirac equation (© Springer Verlag, from Ref. 172).

This is the Dirac equation in the (less usual) spinorial representation.* The last term shows that mass mixes right and left chiralities. The equation can be expanded to include potentials using minimal coupling, in the same way as done above for the Schrödinger

* The matrix γ_5 is defined here as

$$\gamma_5 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} , \tag{143}$$

where I is the 2×2 identity matrix.

and Pauli equations.

The above textbook derivation of the Dirac equation from usual quantum theory can be repeated and visualized also with the help of strands. There is no difference in arguments or results. The derivation with the help of strands was performed for the first time by Battey-Pratt and Racey, in 1980. They explored a central object connected by unobservable strands (or 'tails') to the border of space, as shown in Figure 45. In their approach, the central object plus the tails correspond to a quantum particle. The central object is assumed to be continuously rotating, thus reproducing spin 1/2. They also assumed that only the central object is observable. (In the strand model, the central object becomes the tangle core.) Battey-Pratt and Racey then explored a relativistically moving object of either chirality. They showed that a description of such an object requires four complex fields. Studying the evolution of the phases and axes for the chiral objects yields the Dirac equation. The derivation by Battey-Pratt and Racey is mathematically equivalent to the textbook derivation just given.

Ref. 172

We can thus say that the Dirac equation follows from the belt trick. We will visualize this connection in more detail in the next section. When the present author found this connection in 2008, Lou Kauffman pointed out the much earlier paper by Battey-Pratt and Racey. In fact, Paul Dirac was still alive when they found this connection, but unfortunately he did not answer their letter asking for comment.

In summary, tangles completely reproduce both the rotation and the linear motion of elementary fermions. Therefore, the strand model provides a simple view on the evolution equations of quantum theory. In the terms of the strand model, when spin is neglected, the Schrödinger equation describes the evolution of crossing density. For relativistic fermions, when the belt trick is included, the Dirac equation describes the evolution of crossing density. In fact, strands visualize these evolution equations in the most concrete way known so far.

Visualizing spinors and Dirac's equation using tangles

Despite its apparent complexity, the Dirac equation makes only a few statements: spin 1/2 particles are fermions, obey the relativistic energy–momentum relation, keep the quantum of action invariant, and thus behave like a wave. Each statement is visualized by the tangle model of fermions: tangles behave as spinors, the relativistic energy–momentum relation is built-in, the fundamental principle holds, and rotating tangle cores reproduce the evolution of the phase. Let us look at the details.

Given a particle tangle, the short-time fluctuations lead, after averaging of the crossings, to the wave function. The tangle model of fermions also provides a *visualization* of the *spinor* wave function. Indeed, at each point in space, the wave function has the following parameters:

— There is an average density $\rho(x, t)$; physically, this is the probability density. In the strand model, this is the local crossing density.
— There is a set of three Euler angles α, β and γ; physically, they describe the average local orientation and phase of the spin axis. In the strand model, this is the average local orientation and phase of the tangle core.
— There is a second set of three parameters $\boldsymbol{v} = (v_x, v_y, v_z)$; physically, they describe, at one's preference, either the average local Lorentz boost or a second set of three Euler

angles. In the strand model, these parameters describe the average local deformation of the core that is due to the Lorentz boost. It can also be seen as the axis around which the belt trick is performed.

— There is a phase δ; physically, this represents the relative importance of particle and antiparticle density. In the strand model, this phase describes with what probability the average local belt trick is performed right-handedly or left-handedly.

In total, these are eight real parameters; they correspond to one positive real number and seven phases. They lead to the description of a spinor wave function as
Ref. 173

$$\varphi = \sqrt{\rho}\, e^{i\delta}\, L(\boldsymbol{v})\, R(\alpha/2, \beta/2, \gamma/2) \;, \tag{144}$$

where the product LR is an abbreviation for the boosted and rotated unit spinor and all parameters depend on space and time. This expression is equivalent to the description with four complex parameters used in most textbooks. In fact, this description of a spinor wave function and the related physical visualization of its density and its first six phases dates already from the 1960s. The visualisation can be deduced from the study of relativistic spinning tops or of relativistic fluids. Rotating tangles are more realistic, however. In contrast to all previous visualizations, the rotating tangle model explains also the last, seventh phase. This is the phase that describes matter and anti-matter, that explains the appearance of the quantum of action $\hbar$, and that explains the fermion behaviour.
Ref. 173

In short, only rotating tangles together with the fundamental principle provide a simple, complete and precise visualisation of spinor wave functions and their evolution. The tangle model for spinning relativistic quantum particles remains a simple extension of Feynman's idea to describe a quantum particle as a rotating little arrow. The arrow can be imagined as being attached to the rotating tangle core. The tails are needed to reproduce fermion behaviour. The specific type of tangle core determines the type of particle. The blurring of the crossings defines the wave function. Rotating arrows describe non-relativistic quantum physics; rotating tangles describe relativistic quantum physics.
Ref. 165

Visualizing spinor wave functions with tangles of strands helps the understanding of the Dirac equation in several ways.

1. Tangles support the view that elementary particles are little rotating entities, also in the relativistic case. This fact has been pointed out by many scholars over the years. The strand model provides a consistent visualization for these discussions. Ref. 173
2. The belt trick can be seen as the mechanism underlying the famous Zitterbewegung that is part of the Dirac equation. The limitations in the observing the belt trick translate directly into the difficulties of observing the Zitterbewegung. Ref. 174
3. The belt trick also visualizes why the velocity operator for a relativistic particle has eigenvalues $\pm c$.
4. The Compton length is often seen as the typical length at which quantum field effects take place. In the tangle model, it would correspond to the average size needed for the belt trick. The strand model thus suggests that the mass of a particle is related to the average size needed for the belt trick.
5. Tangles support the – at first sight bizarre – picture of elementary particles as little charges rotating around a centre of mass. Indeed, in the tangle model, particle rota- Ref. 175

Page 177 tion requires a regular application of the belt trick of Figure 19, and the belt trick can be interpreted as inducing the rotation of a charge, defined by the tangle core, around a centre of mass, defined by the average of the core position. It can thus be helpful to use the strand model to visualize this description.

6. The tangle model can be seen as a vindication of the stochastic quantization research programme; quantum motion is the result of underlying fluctuations. For example, the similarity of the Schrödinger equation and the diffusion equation is modelled and explained by the strand model: since crossings can be rotated, diffusion of crossings leads to the imaginary unit that appears in the Schrödinger equation. Ref. 176

In short, rotating tangles are a correct underlying model for the propagation of fermions. And so far, tangles are also the only known correct model. *Tangles model propagators.* This modelling is possible because the Dirac equation results from only three ingredients:

— the relation between the quantum of action and the phase of the wave function (the wave behaviour),
— the relation between the quantum of action and spinor behaviour (the exchange behaviour),
— and the mass–energy relation of special relativity (the particle behaviour), itself due to the fundamental principle.

And all three ingredients are reproduced by the strand model. We see that the apparent complexity of the Dirac equation hides its fundamental simplicity. The strand model reproduces the ingredients of the Dirac equation, reproduces the equation itself, and makes the simplicity manifest. In fact, we can say:

▷ The Dirac equation describes the relativistic infinitesimal belt trick or string trick.

The belt trick is fundamental for understanding the Dirac equation. In the strand model, core rotations vary along two dimensions – the rotation is described by two angles – and so does the belt trick. The resulting four combinations form the four components of the Dirac spinor and of the Dirac equation. Page 180

In summary, tangles can be used as a precise visualization and explanation of quantum physics. Wave functions, also those of fermions, are *blurred tangles* – with the detail that not the strands, but their crossings are blurred.

Quantum mechanics vs. quantum field theory

Quantum mechanics is the approximation to quantum physics in which fields are continuous and particles are immutable. In the strand model, quantum mechanics is thus the approximation in which a particle is described by a tangle with a shape that is *fixed* in time. This approximation allows us to derive the Dirac equation, the Klein–Gordon equation, the Proca equation, the Pauli equation and the Schrödinger equation. In this approximation, the strand model for the electron in a hydrogen atom is illustrated in Figure 46. This approximation already will allow us to deduce the existence of the three gauge interactions, as we will see in the next chapter.

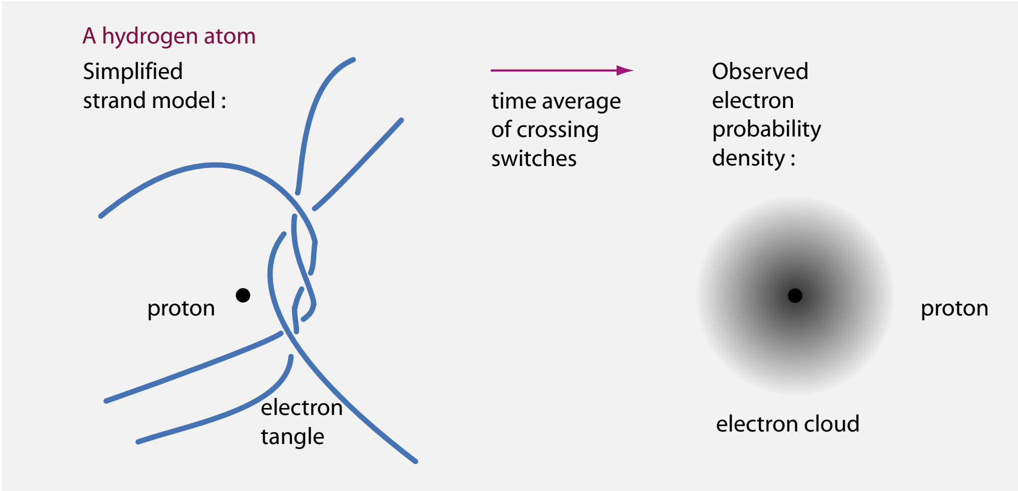


FIGURE 46 A simple, quantum-mechanical view of a hydrogen atom.

In contrast, *quantum field theory* is the description in which fields are themselves described by bosons, and particles types can transform into each other. The strand model allows us to deduce the existence of all known gauge bosons, as shown in the next chapter. In the strand description of quantum field theory, particles are not tangles with a fixed shape of their core, but for each particle, the shape *varies*. This variation leads to gauge boson emission and absorption.

A flashback: settling three paradoxes of Galilean physics

In all descriptions of physics, space and time are measured, explained and defined using matter. This occurs, for example, with the help of metre bars and clocks. On the other hand, matter is measured, explained and defined using space and time. This occurs, for example, by following a localized body over space and time. The circularity of the two definitions is at the basis of modern physics.

Page 110
As already mentioned above, the circularity is a natural consequence of the strand model. Both matter and space-time turn out to be approximations of the same basic building blocks; this common origin explains the apparent circular reasoning of Galilean physics. Most of all, the strand model changes it from a paradox to a logical necessity.

The strand model defines vacuum, and thus physical space, as a result of averaging strand crossings. Space is thus a *relative* concept. Newton's bucket experiment is sometimes seen as a counter-argument to this conclusion and as an argument for absolute space. However, the strand model shows that any turning object is connected to the rest of the universe through its tails. This connection makes every rotation an example of relative motion. Rotation is thus always performed relatively to the horizon of the universe. On the other hand, the detection of tangles among the tails allows a *local* determination of the rotation state, as is observed. Strands thus confirm that rotation and space are relative concepts. Strands thus also explain why we can turn ourselves on ice by rotating an arm over our head, without outside help. Strands lie to rest all issues around the rotating bucket.

A long time ago, Zeno of Elea based one of his paradoxes – the flying arrow that cannot reach the target – on an assumption that is usually taken as granted: he stated the impossibility to distinguish a short-time image (or state) of a *moving* body from the image (or state) of a *resting* body. The flattening of the tangles involved shows that the assumption is incorrect; motion and rest are *distinguishable*, even in (imagined) photographs taken with extremely short shutter times. The argument of Zeno is thus not possible, and the paradox disappears.

FUN CHALLENGES ABOUT QUANTUM THEORY

> Urlaub ist die Fortsetzung des Familienlebens unter erschwerten Bedingungen.*
> Dieter Hildebrandt

Challenge 143 s Are the definitions for the addition and multiplication of Schrödinger wave functions that were given above also valid for spinor tangle functions?

* *

Challenge 144 e The definition of tangle functions, or wave functions, did not take into account the crossings of the vacuum strands, but only those of the particle tangle. Why is this allowed?

* *

Challenge 145 e Modelling the measurement of action at the quantum level as the counting of full turns of a wheel is a well-known idea that is used by good teachers to take the mystery out of quantum physics. The strand model visualizes this idea by assigning the quantum of action $\hbar$ to a full turn of one strand segment around another.

* *

Challenge 146 s Is any axiomatic system of quantum theory in contrast with the strand model?

* *

Challenge 147 s In the strand model, tangle energy is related to tangle core rotation. What is the difference between the angular frequency for tangles in the non-relativistic and in the relativistic case?

* *

Ref. 177 Challenge 148 e If you do not like the deduction of quantum mechanics given here, there is an alternative: you can deduce quantum mechanics in the way Schwinger did in his course, using the quantum action principle.

* *

Ref. 178 Challenge 149 r Modern teaching of the Dirac equation replaces the spinor picture with the vector picture. Hrvoje Nikolić showed that the vector picture significantly simplifies the understanding of Lorentz covariance of the Dirac equation. How does the vector picture clarify the relation between the belt trick and the Dirac equation?

* 'Vacation is the continuation of family life under aggravated conditions.' Dieter Hildebrandt (b. 1927 Bunzlau, d. 2013 Munich) was a cabaret artist, actor and author.

* *

In the strand description of quantum mechanics, strands are impenetrable: they cannot pass through each other (at finite distances). Can quantum mechanics also be derived if Challenge 150 s the model is changed and this process is allowed? Is entanglement still found?

* *

A puzzle: Is the belt trick possible in a continuous and deformable medium – such as a Challenge 151 e sheet or a mattress – in which a coloured sphere is suspended? Is the belt trick possible with an *uncountably* infinite number of tails?

* *

Page 181 At first sight, the apheresis machine diagram of Figure 24 suggests that, using the belt trick, animals could grow and use wheels instead of legs, because rotating wheels could be supplied with blood and connected to nerves. Why did wheels not evolve neverthe- Challenge 152 s less?

SUMMARY ON QUANTUM THEORY OF MATTER: EXPERIMENTAL PREDICTIONS

In this chapter, we used the fundamental principle – crossing switches define the quantum of action $\hbar$ and the other Planck units – to deduce that particles are tangles of strands and that wave functions are time-averaged rotating tangles. In simple words,

▷ Both non-relativistic and relativistic wave functions are *blurred rotating tangles*.

More precisely, a wave function appears from the blurred crossings of a tangle. The components and phases of the wave function at a point in space are due to the orientation and phase of crossings at that point. We also deduced that blurred tangles obey the least action principle and the Dirac equation.

In other words, visualizing the quantum of action as a crossing switch implies quantum theory. The strand model confirms Bohr's statement: quantum theory is indeed a consequence of the quantum of action. Specifically, the strand model thus shows that all quantum effects are *consequences of extension* and *consequences of the three dimensions of space*. More precisely, all quantum effects are *due to tails*, the tails of the tangles that represent a quantum system. In particular, the strand model confirms that

▷ The Dirac equation is essentially the infinitesimal version of the belt trick (or string trick).

In other words, strands also reproduce also the *propagator* of quantum particles.

As a result, we have shown that strands reproduce the relativistic Lagrangian density

$\mathcal{L}$ of charged, elementary, relativistic fermions in an external electromagnetic field $\boldsymbol{A}$

$$\mathcal{L} = \overline{\varphi}\,(i\hbar c \not{D} - c^2 m)\,\varphi\ , \tag{145}$$

where

$$\not{D} = \gamma^\sigma \mathrm{D}_\sigma = \gamma^\sigma(\partial_\sigma - iqA_\sigma)\ . \tag{146}$$

We thus conclude that *strands reproduce the quantum theory of matter*.

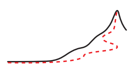
The strand model predicts deviations from the relativistic matter Lagrangian, and thus from the Dirac equation, *only* in three cases: first, when quantum aspects of electrodynamic field play a role, second, when nuclear interactions play a role, and third, when space curvature, i.e., strong gravity, plays a role. All this agrees with observation.

We will deduce the description of quantum electrodynamics and of the nuclear interactions in the next chapter. In the case of gravity, the strand model predicts that deviations from quantum theory occur exclusively when the energy–momentum of an elementary particle approaches the Planck value, i.e., for really strong gravity. Such deviations are not accessible to experiment at present. We will explore this situation in the subsequent chapter.

In addition, the strand model predicts that in nature, the Planck values for momentum Page 40 and energy are limit values that cannot be exceeded by a quantum particle. All experiments agree with this prediction.

The deduction of quantum theory from strands given here is, at present, the *only* known microscopic explanation for quantum physics. So far, no other microscopic model, no different explanation nor any other Planck-scale deduction of quantum theory has been found. In particular, the extension of fundamental entities – together with observability limited to crossing switches – is the key to understanding quantum physics.

Page 164 Let us evaluate the situation. In our quest to explain the open issues of the millennium list, we have explained the origin of Planck units, the origin of wave functions, the origin of the least action principle, the origin of space-time dimensions, the Lorentz and Poincaré symmetries, the origin of particle identity, and the simplest part of the Lagrangian of quantum field theory, namely, the Lagrangian of free fermions, such as the electron, and that of fermions in continuous external fields. Therefore, for the next leg, we turn to the most important parts of the standard model Lagrangian that are missing: those due to gauge interactions.



Chapter 9

GAUGE INTERACTIONS DEDUCED FROM STRANDS

What are interactions? At the start of this volume, when we summarized what relates the Planck units to relativity and to quantum theory, we pointed out that the nature of interactions at Planck scales was still in the dark. In the year 2000, it was known for several decades that the essential properties of the electromagnetic, the weak and the strong nuclear interaction are their respective gauge symmetries: all three interactions are *gauge interactions*. But the underlying reason for this property was still unknown.

Page 18

Ref. 179

In this chapter we discover that fluctuating strands in three spatial dimensions explain the existence of precisely three gauge interactions, each with precisely the gauge symmetry group that is observed. This is the first time ever that such an explanation is possible. In other terms, we will deduce quantum field theory from strands. Indeed, strands provide a natural mechanism for interactions that explains and implies Feynman diagrams. The term 'mechanism' has to be taken with a grain of salt, because there is nothing mechanical involved; nevertheless, the term is not wrong, because we shall discover a surprisingly simple result: *Gauge interactions and gauge symmetries are due to specific strand deformations.*

Ref. 180

In this chapter, we work in *flat* space-time, as is always done in quantum field theory. We leave the quantum aspects of *curved* space-time and of gravitation for the next chapter. We thus start by exploring the non-gravitational interactions in the quantum domain.

Interactions and phase change

Experiments in the quantum domain show that interactions *change the phase* of wave functions. But how precisely does this happen? The strand model will give us a simple answer: the emission and the absorption of gauge bosons is only possible *together* with a phase change. To explain this connection, we need to study the phase of tangle *cores* in more detail.

When we explored spin and its connection to the belt trick, we pictured the rotation of the tangle core in the same way as the rotation of a belt buckle: we assumed that the core of the tangle rotates like a *rigid* object. The rotation is achieved through the shape fluctuations of the tails only. Why did we assume this?

Page 176

In Feynman's description of quantum theory, *free particles are advancing rotating arrows.* In the strand model, *free* particle motion is modelled as the change of position of the tangle core and *spin* as the rotation of the core. We boldly assumed that the core

Ref. 165

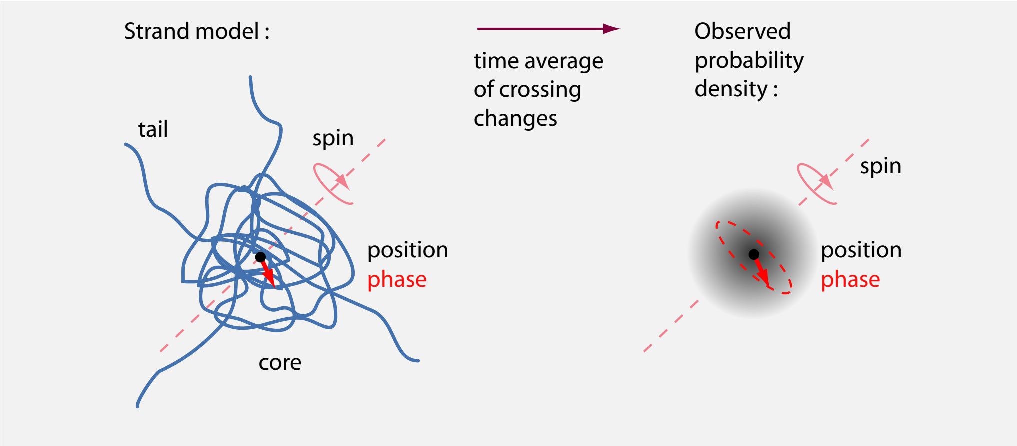


FIGURE 47 In the chapter on quantum theory, the phase was defined assuming a *rigidly rotating core*; this approximation was also used in the description of particle translation.

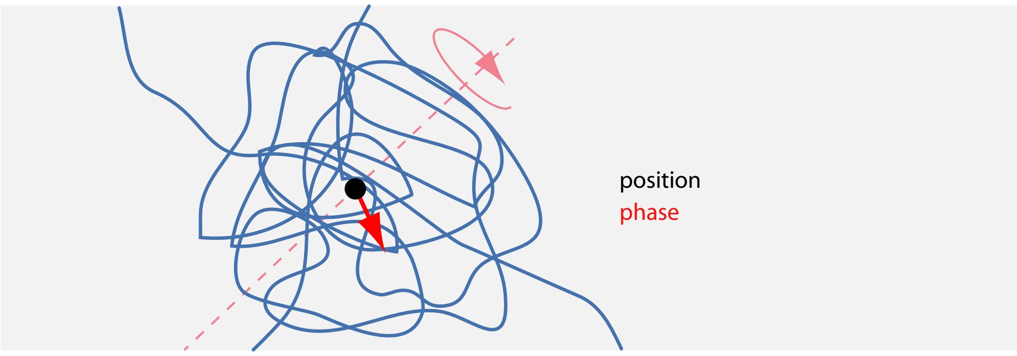


FIGURE 48 A magnified tangle core shows that the phase can also change due to *core deformations*; such core deformations lead to *gauge interactions*.

remained rigid, attached the phase arrow to it, and described spin as the rotation of the core with its attached arrow, as shown again in Figure 47. This bold simplification led us to the Dirac equation. In short, the assumption of a rigid core works. But what happens if the core is *not* rigid?

We know from observation and from quantum theory that

▷ An ***interaction*** is a process that changes the phase of a wave function, but differs from a rotation.

In the strand model, shape deformations of tangle cores also lead to phase changes – and such deformations differ from a rotation. In fact, we will discover that core deformations automatically lead to precisely those three gauge interactions that we observe in nature.

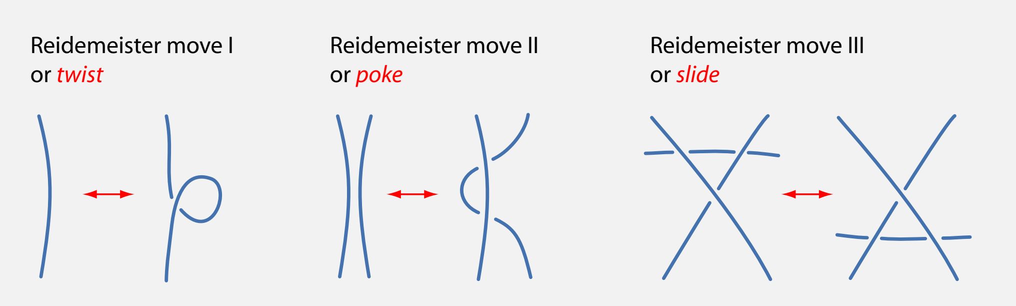


FIGURE 49 The Reidemeister moves: the three types of deformations that induce crossing switches – if the moves are properly defined in three dimensions.

Tail deformations versus core deformations

We can summarize the previous chapter, on the free motion of matter tangles, as the chapter that focused on shape fluctuations of *tails*. Indeed, the belt trick completed the proof that

▷ ***Space-time symmetries*** are due to *tail* deformations.

All space-time symmetries – translation, rotation, boost, spin and particle exchange – are due to tail deformations; in such tail deformations, the tangle core is assumed to remain unchanged and rigid (in its own rest frame).

In contrast, the present chapter focuses on shape fluctuations in *tangle cores*.We will discover that

▷ ***Gauge symmetries*** are due to *core* deformations.

Let us explore the tangle core in more detail. Figure 48 shows a magnified view of the core and its phase arrow. The phase of the core results from the phases of all its crossings. The figure illustrates that the phase arrow will be sensitive to the shape fluctuations and deformations of the strand segments that make up the core.

In nature, any phase change of the wave function that is not due to a space-time symmetry is due to an interaction. For the strand model, this connection implies:

▷ When the phase of a core changes through *rigid orientation change*, we speak of ***core rotation***.

▷ When the phase of a core changes through *core shape deformation*, we speak of ***interaction***.

We thus need to understand two things: First, what kinds of core deformation exist? Secondly, how precisely is the phase – i.e., each arrow definition – influenced by core deformations? In particular, we have to check the answers and deductions with experiment.

The first question, on the classification of the core deformations, is less hard than

Ref. 180 it might appear. The fundamental principle – events are crossing switches of strands – implies that deformations are observable only if they induce crossing switches. Other deformations do not have any physical effect. (Of course, certain deformations will have crossing switches for one observer and none for another. We will take this fact into consideration.) Already in 1926, the mathematician Kurt Reidemeister classified all those Ref. 182 tangle deformations that lead to crossing switches. The classification yields exactly three classes of deformations, today called the three *Reidemeister moves*. They are shown in Figure 49.

▷ The ***first Reidemeister move***, or ***type I move***, or ***twist***, is the addition or removal of a twist in a strand.

▷ The ***second Reidemeister move***, or ***type II move***, or ***poke***, is the addition or removal of a bend of one strand under (or over) a second strand.

▷ The ***third Reidemeister move***, or ***type III move***, or ***slide***, is the displacement of one strand segment under (or over) the crossing of two other strands.

The type number of each Reidemeister move is also the number of involved strands. We will discover that despite appearances, each Reidemeister move induces a crossing switch. To find this connection, we have to generalize the original Reidemeister moves, which were defined in a two-dimensional projection plane, to the three-dimensional situation of tangle cores.

The three Reidemeister moves turn out to be related to the three gauge interactions:

▷ The first Reidemeister move corresponds to *electromagnetism*. The second Reidemeister move corresponds to the *weak nuclear interaction*. The third Reidemeister move corresponds to the *strong nuclear interaction*.

We will prove this correspondence in the following.

For each Reidemeister move we will explore two types of core deformation processes: One deformation type are *core fluctuations*, which correspond, as we will see, to the emission and absorption of *virtual* interaction bosons. The other deformations are *externally induced core disturbances*, which correspond to the emission and absorption of *real* interaction bosons. As the first step, we show that both for fluctuations and for disturbances, the first Reidemeister move, the twist, is related to the electromagnetic interaction.

Electrodynamics and the first Reidemeister move

Experiments show that electromagnetism is described by potentials. Experiments also show that potentials change the phase, the rotation frequency and the wave number of wave functions. In particular, for electromagnetism, the potentials are due to the flow of real and virtual, massless, uncharged spin-1 photons. Photons are emitted from or absorbed by charged elementary particles; neutral elementary particles do not emit or absorb photons. There are two types of electric charge, positive and negative. The attrac-

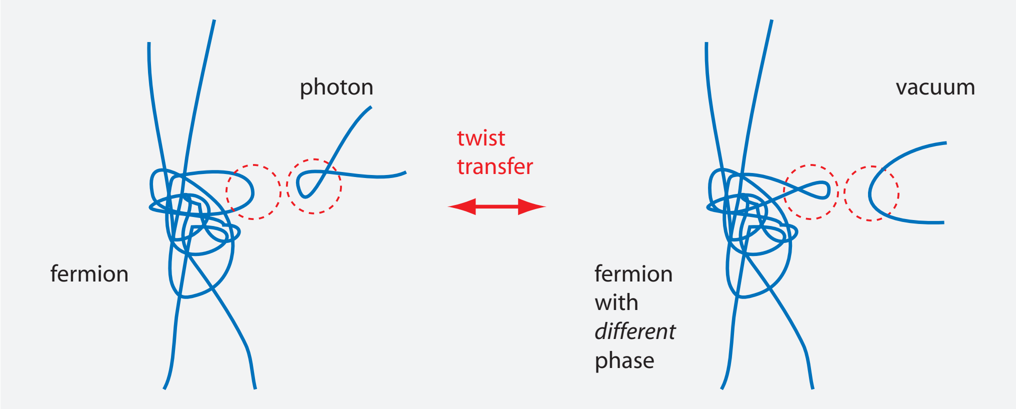


FIGURE 50 A single strand changes the rotation of a tangle: *twist transfer* is the basis of electromagnetism in the strand model. No strand is cut or reglued; the transfer occurs statistically, through the excluded volume due to the impenetrability of strands. Twist transfer generates a U(1) gauge group, as explained in the text.

tion and repulsion of static charges diminishes with the inverse square of the distance. Charge is conserved. All charged particles are massive and move slower than light. The Lagrangian of matter coupled to the electromagnetic field has a U(1) gauge symmetry – it is described by minimal coupling. Electromagnetism has a single fundamental Feynman diagram. The electromagnetic coupling constant at low energy, the so-called *fine structure constant*, is measured to be $\alpha = 1/137.035\,999\,139(31)$; its energy dependence is described by renormalization.

Ref. 5

The previous paragraph contains everything known about the electromagnetic interaction. For example, Maxwell's field equations follow from Coulomb's inverse square relation, its relativistic generalization, and the conservation of charge. More precisely, all experimental observations about electricity and magnetism follow from the Lagrangian of quantum electrodynamics, or QED. In short, we now need to show that the Lagrangian of QED follows from the strand model.

Ref. 183

Strands and the twist, the first Reidemeister move

In the strand model of electromagnetism, massless spin 1 bosons such as the photon are made of a single strand. How can a single strand change the phase of a tangle? The answer is given in Figure 50: a *twisted loop* in a single strand will influence the rotation of a tangle because it changes the possible shape fluctuations of the tangle core. Due to the impenetrability of strands, an approaching twisted loop will sometimes transfer its twist to the tangle: this process will deform the tangle core and thereby change its phase. The observed effect of an electromagnetic field on the phase of a charged fermion is the *time average* of all such twist transfers.

Page 176

Single strands represent bosons, as we saw above. Rotating the core of the twist around the tethers by 2π gives back the original configuration: twists have spin 1. Twisted loops are single strands and can have *two* twist senses, or two polarizations. Single, twisted and *unknotted* strands have no mass; in other words, twisted loops effectively move with the speed of light. And twisted loops, being curved, carry energy.

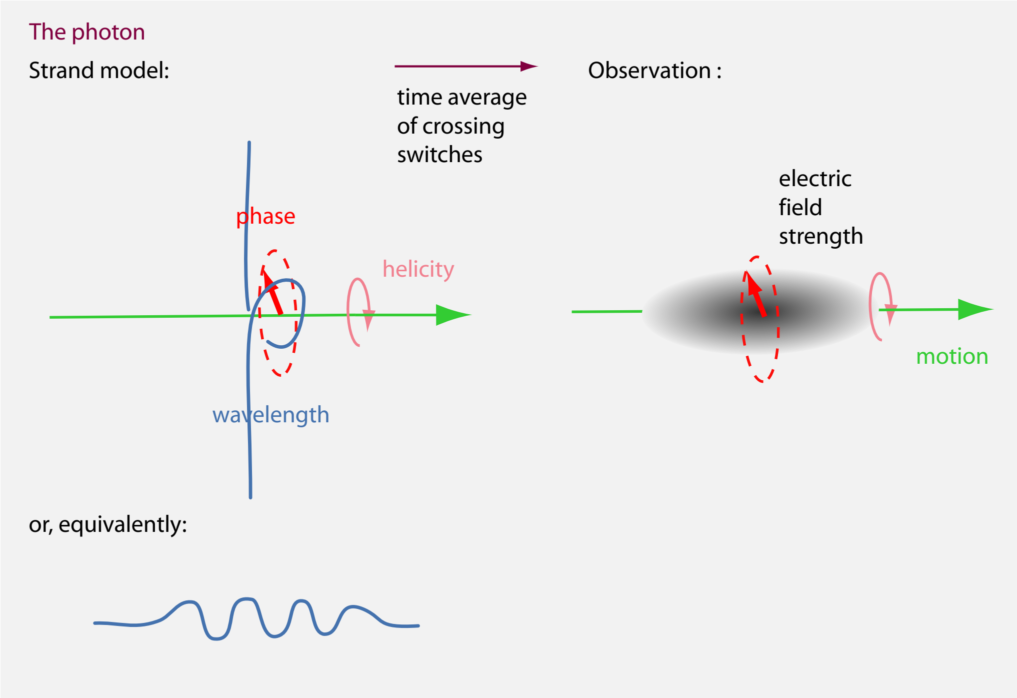


FIGURE 51 The photon in the strand model.

Approaching twisted loops will change the phase, i.e., the orientation of a matter tangle. Twisted loops correspond to a local rotation of a strand segment by π. But twists can be generalized to arbitrary angles. These generalized twists can be concatenated. Because they are described by a single angle, and because a double twist is equivalent to no twist at all, twists form a U(1) group. We show this in detail shortly. Page 235

In summary, twists behave like *photons* in all their properties. Therefore, the strand model suggests:

▷ A ***photon*** is a twisted strand. An illustration is given in Figure 51.

▷ The ***electromagnetic interaction*** is the transfer of twists, i.e., the transfer of first Reidemeister moves, between two particles, as shown in Figure 50.

The transfer of a twist from a single strand to a tangle core thus models the absorption of a photon. We stress again that this transfer results from the way that strands hinder each other's motion, because of their impenetrability. No strand is ever cut or reglued.

Can photons decay, disappear or break up?

The strand model of the photon, as shown in Figure 51, might be seen to suggest that photons can disappear. For example, if a photon strand is straightened out by pulling the ends of the helical deformation, the helix might disappear. A helix might also disappear by a shape fluctuation or transform into several helices. However, this is a fallacy.

A lone twist cannot disappear by pulling; ''pulling'' requires an apparatus that performs it. That is impossible. A lone twist cannot disappear by fluctuations either, because a photon also includes the vacuum strands around it. In the strand model, the energy of the photon is localized in the configuration formed by the photon strand and the surrounding vacuum strands. In the strand model, energy is localized in regions of strand curvature. If the helical strands disappears, the surrounding vacuum strands are curved instead, or more strongly, and the twist energy is taken up by these surrounding strands. The net result is that the helix is transferred, permanently or for a short time, to another strand. In other terms, in the strand model, photons can also move by hopping from one strand to the next.

Page 351

Also, a single photon strand cannot break up into *several* photon strands of smaller helical diameters or of different rotation frequencies. Such a process is prevented by the fundamental principle, when the vacuum is taken into account.

Challenge 153 e

The only way in which a photon can disappear completely is by transferring its crossing, i.e., its energy to a tangle. Such a process is called the *absorption* of a photon by a charged particle.

In short, due to energy and to topological restrictions, the strand model prevents the decay, disappearance or splitting of photons, as long as no electric charge is involved. Linear and angular momentum conservation also lead to the same conclusion. Photons are *stable* particles in the strand model.

Electric charge

Surrounded by a bath of photon strands, not all fermion tangles will change their phase. A tangle subject to randomly approaching virtual photons will feel a net effect over time only if it lacks some symmetry. In other words, only tangles that lack a certain symmetry will be electrically charged. Which symmetry will this be?

In a bath of photon strands, thus in a bath that induces random Reidemeister I moves, only *chiral* fermion tangles are expected to be influenced. In other terms:

▷ ***Electric charge*** is due to lack of mirror symmetry, i.e., to tangle chirality.

Conversely, we have:

▷ ***Electrically charged particles*** randomly emit twisted strands. Due to the tangle chirality, a random emission will lead to a slight asymmetry, so that right-handed twists will be in the majority for particles of one charge, and left-handed twists will be in the majority for particles of the opposite charge.

Equating electric charge with tangle chirality allows modelling several important observations. First, because chirality can be right-handed or left-handed, there are positive and negative charges. Second, because strands are never cut or reglued in the strand model, chirality, and thus electric charge, is a *conserved quantity*. Third, chirality is only possible for tangles that are localized, and thus massive. Therefore, chiral tangles – charged particles – always move slower than light. Fourth, a chiral tangle at rest induces a twisted strand density around it that changes as $1/r^2$, as is illustrated in Figure 52. Finally,

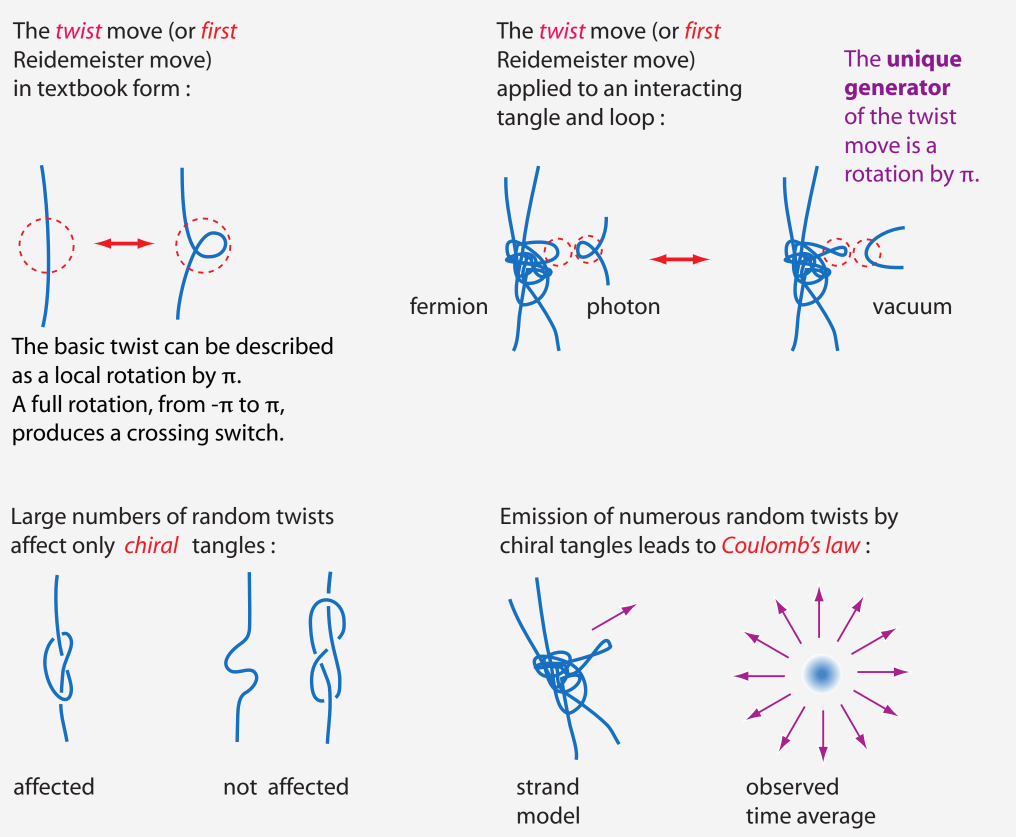


FIGURE 52 Electromagnetism in the strand model: the electromagnetic interaction, electric charge and Coulomb's inverse square relation. (This image needs to be updated: no knotted tangles occur in nature.)

photons are uncharged; thus they are not influenced by other photons (to first order).

In short, all properties of electric charge found in nature are reproduced by the tangle model. We now check this in more detail.

Challenge: What topological invariant is electric charge?

Chirality explains the sign of electric charge, but not its magnitude in units of the elementary charge e. A full definition of electric charge must include this aspect.

Mathematicians defined various topological invariants for knot and tangles. *Topological invariants* are properties that are independent of the shape of the knot or tangle, but allow to distinguish knots or tangles that differ in the ways they are knotted or tangled up. Several invariants are candidates as building blocks for electric charge: *chirality* c, which can be $+1$ or -1, *minimal crossing number* n, or *topological writhe* w, i.e., the signed minimal crossing number.

Page 437

A definition of electric charge q, proposed by Claus Ernst, is $q = c\,(n \bmod 2)$. Another option for the definition of charge is $q = w/3$. Equivalent definitions use the linking number. At this point of our exploration, the issue is open. We will come back to the

Page 390 detailed connection between charge, chirality and tangle topology later on.

Electric and magnetic fields and potentials

The definition of photons with twisted strands leads to the following definition.

▷ The ***electric field*** is the volume density of (oriented) crossings of twisted loops.

▷ The ***magnetic field*** is the flow density of (oriented) crossings of twisted loops.

▷ The ***electric potential*** is the density of twisted loops.

▷ The ***magnetic potential*** is the flow density of twisted loops.

The simplest way to check these definitions is to note that the random emission of twisted loops by electric charges yields Coulomb's inverse square relation: the force between two static spherical charges changes with inverse square of the distance. The strand model implies that in this case, *the crossing density is proportional to the square of the loop density*; in other words, the potential falls of as the inverse distance, and the electric field as the square distance.

The definition of the magnetic field simply follows from that of the electric field by changing to moving frame of reference. The two field definitions are illustrated in Figure 53.

We note that the electric field is defined almost in the same way as the wave function: Page 186 both are oriented crossing densities. However, the electric field is defined with the crossing density of *twisted loops*, whereas the wave function is defined with the crossing density of *tangles*. The definitions differ only by the topology of the underlying strand structures.

In the strand model, energy, or action per time, is the number of crossing switches *per time*. The electromagnetic field energy per volume is thus given by the density of crossing switches *per time* that are due to twisted loops. Now, the strand model implies that *the crossing switch density per time is given by half the square of the crossing density* Challenge 154 e *plus half the square of the crossing density flow*. For twisted loops, we thus get that the energy density is half the square of the electric plus half the square of the magnetic field. Inserting the proportionality factors that lead from Planck units to SI units we get the well-known expression

$$\frac{E}{V} = \frac{\varepsilon_0}{2}E^2 + \frac{1}{2\mu_0}B^2 \; . \qquad (147)$$

The strand model thus reproduces electromagnetic energy.

We note that in the strand model, the definition of the fields implies that there is no *magnetic charge* in nature. This agrees with observation.

The strand model predicts limit values to all observables. They always appear when strands are as closely packed as possible. This implies a maximum electric field value $E_{max} = c^4/4Ge \approx 1.9 \cdot 10^{62}$ V/m and a maximum magnetic field value $B_{max} = c^3/4Ge \approx 6.3 \cdot 10^{53}$ T. All physical systems – including all astrophysical objects, such as gamma-ray

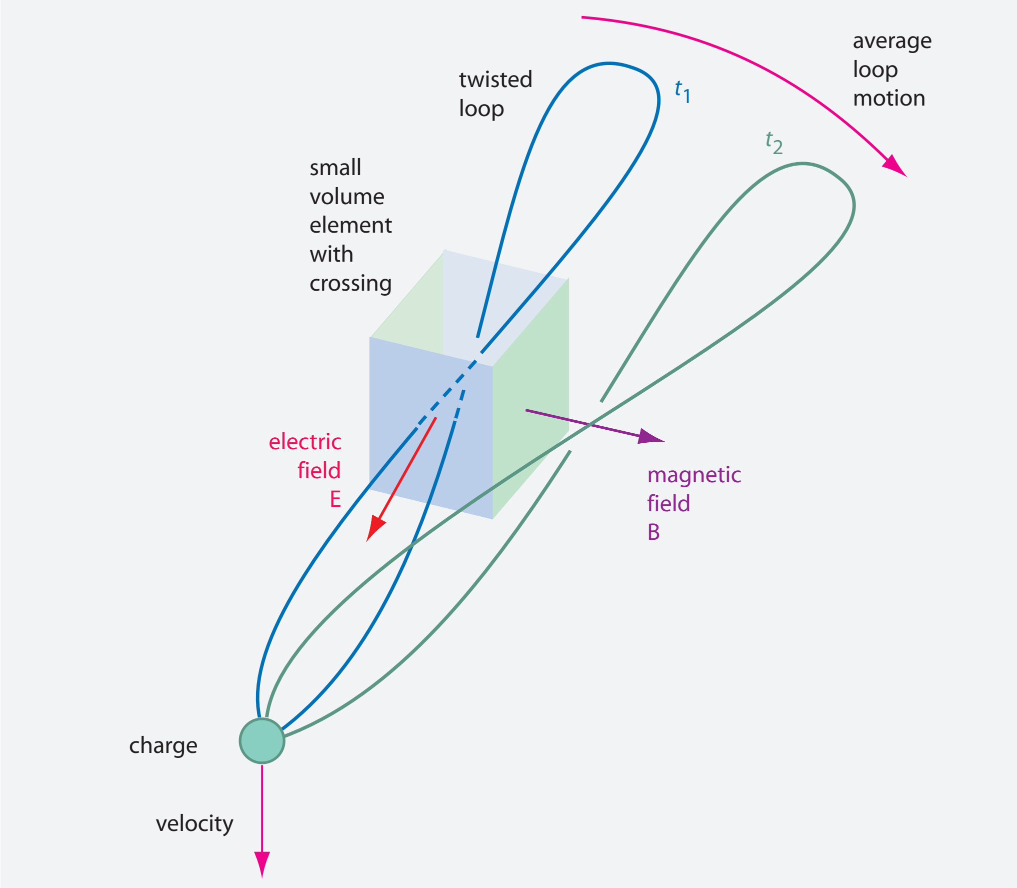


FIGURE 53 Moving twists allow us to define electric fields – as the density of twisted loop crossings – and magnetic fields – as the corresponding flow.

bursters or quasars – are predicted to conform to this limit. This strand model prediction indeed agrees with observations so far.

The Lagrangian of the electromagnetic field

In classical electrodynamics, the energy density of the electromagnetic field is used to deduce its Lagrangian density. The Lagrangian density describes the intrinsic, observer-independent change that occurs in a system. In addition, the Lagrangian density must be quadratic in the fields and be a Lorentz-scalar.

A precise version of these arguments leads to the Lagrangian density of the electromagnetic field F

$$\mathcal{L}_{\mathrm{EM}} = \frac{\varepsilon_0}{2} E^2 - \frac{1}{2\mu_0} B^2 = -\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} \qquad (148)$$

where the electromagnetic field F is defined with the electromagnetic potential A as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \, . \qquad (149)$$

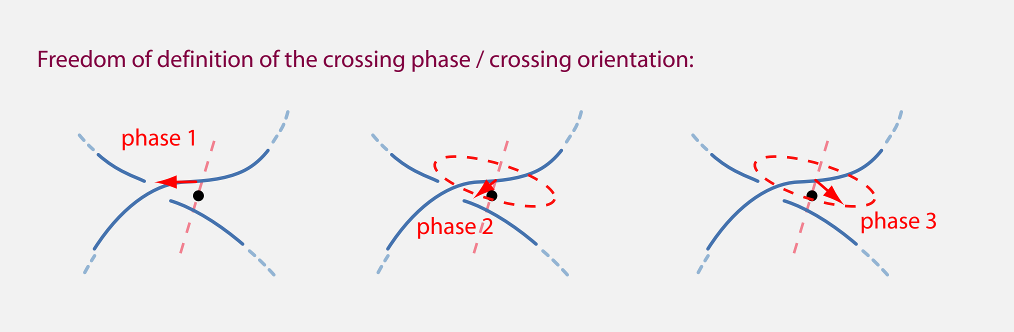


FIGURE 54 The definition of the phase or orientation of a single crossing is not unique: there is a freedom of choice.

Since the strand model reproduces the electromagnetic energy, it also reproduces the Lagrangian of classical electrodynamics. In particular, Maxwell's equations for the electromagnetic field follow from this Lagrangian density. Maxwell's field equations are thus a consequence of the strand model. Obviously, this is no news, because any model that reproduces Coulomb's inverse square distance relation and leaves the speed of light invariant automatically contains Maxwell's field equations. Ref. 183

U(1) GAUGE INVARIANCE INDUCED BY TWISTS

In nature, the electromagnetic potential A_μ is not uniquely defined: one says that there is a freedom in the choice of gauge. The change from one gauge to another is a *gauge transformation*. Gauge transformations are thus transformations of the electromagnetic potential that have no effect on observations. In particular, gauge transformations leave unchanged all field intensities and field energies on the one hand and particle probabilities and particle energies on the other hand.

All these observations can be reproduced with strands. In the strand model, the following definitions are natural:

▷ A ***gauge choice*** for *radiation* and for *matter* is the choice of definition of the respective phase arrow.
▷ A ***gauge transformation*** is a change of definition of the phase arrow.

In the case of electrodynamics, the gauge freedom is a result of allowing phase choices that lie in a plane around the crossing orientation. (The other interactions follow from the other possible phase choices.) The phase choice can be different at every point in space. Changing the (local) phase definition is a (local) gauge transformation. Changing the phase definition for a single crossing implies changing the phase of wave functions and of the electromagnetic potentials. A schematic illustration of the choice of gauge is given in Figure 54 and Figure 55.

We note that gauge transformations have no effect on the density or flow of crossings or crossing switches. In other words, gauge transformations leave electromagnetic field intensities and electromagnetic field energy invariant, as observed. Similarly, gauge transformations have no effect on the number of crossing switches of rotating tangles.

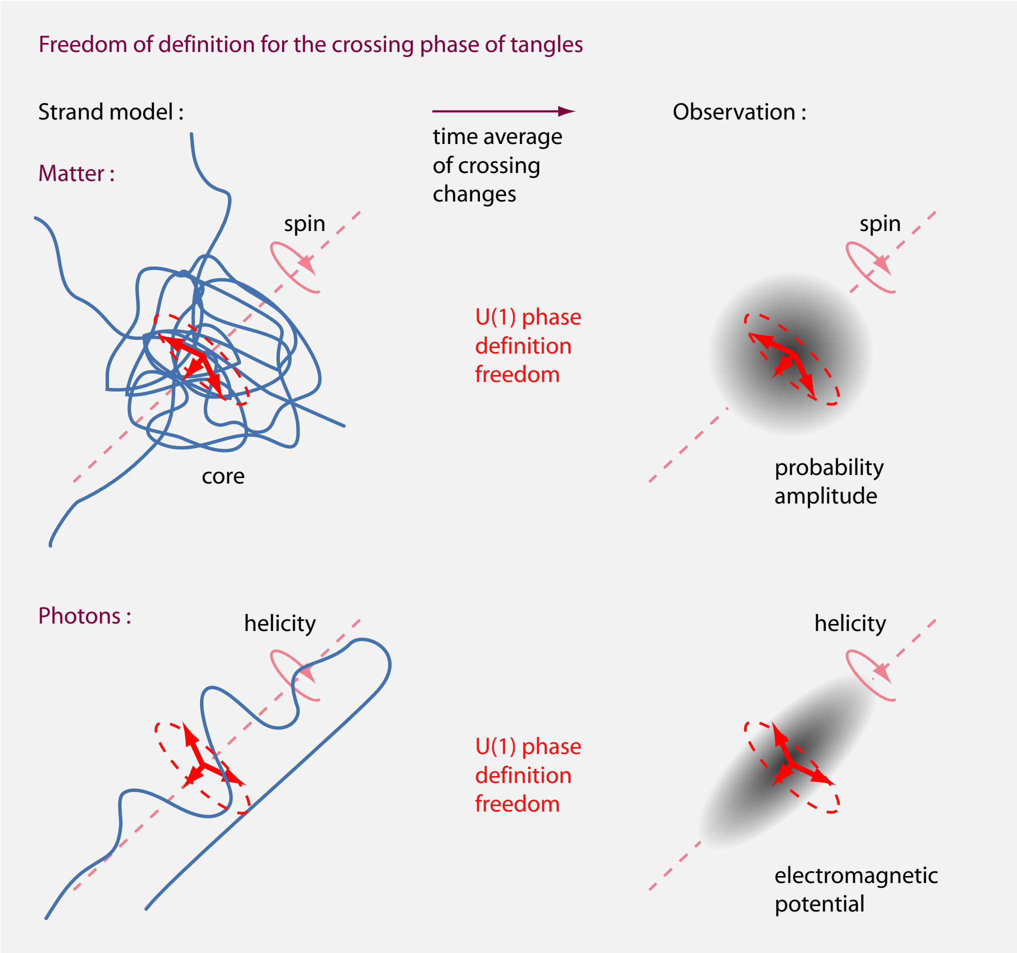


FIGURE 55 The freedom in definition of the phase of crossings leads to the gauge invariance of electrodynamics. Three exemplary choices of phase are shown.

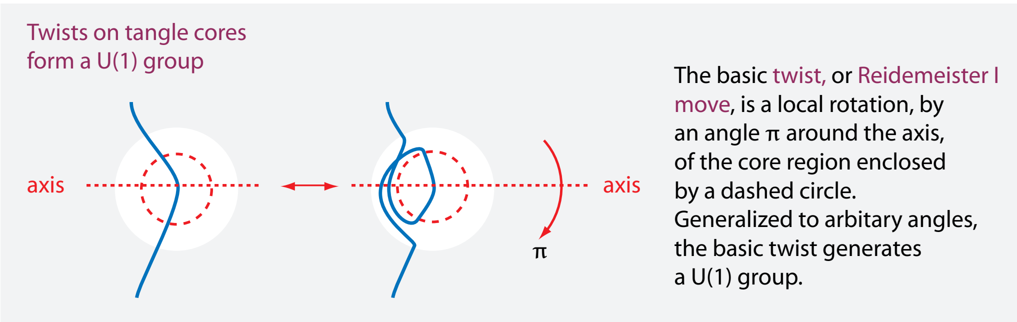


FIGURE 56 How the set of generalized twists – the set of all local rotations of a single strand segment around an axis – forms a U(1) gauge group.

A rotation by 4π does not change the phase, independently of which definition of arrow is chosen. Therefore, gauge transformations leave probability densities – and even observable phase differences – unchanged. This agrees with experiment.

A gauge transformation on a wave functions also implies a gauge transformation on the electrodynamic potential. The strand model thus implies that the two transformations are connected, as is observed. This connection is called *minimal coupling*. In short, minimal coupling is a consequence of the strand model.

U(1) gauge interactions induced by twists

There is only a small step from a gauge *choice* to a gauge *interaction*. We recall:

▷ A ***gauge interaction*** is a change of phase resulting from a strand deformation of the particle core.

In particular, electromagnetism results from the transfer of *twists*; twists are one of the three types of core deformations that lead to a crossing switch.

The basic twist, or first Reidemeister move, corresponds to a local rotation of some strand segment in the core by an angle π, as illustrated by Figure 56. Twists can be generalized to arbitrary angles: we simply define a *generalized twist* as a local rotation of a strand segment by an arbitrary angle. The rotation axis is chosen as shown in Figure 56. Generalized twists can be concatenated, and the identity twist – no local rotation at all – also exists. Generalized twists thus form a group. Furthermore, a generalized twist by 2π is equivalent to no twist at all, as is easily checked with a piece of rope: keeping the centre region is it disappears by pulling the ends, in contrast to a twist by π.

Challenge 155 e

Generalized twists thus behave like $e^{i\vartheta}$. Their concatenation produces a multiplication table

$$\begin{array}{c|c} \cdot & e^{i\pi} \\ \hline e^{i\pi} & 1 \end{array} \qquad (150)$$

that generate and define a U(1) group. In other words, Figure 56 shows that generalized twists define the group U(1), which has the topology of a circle.

In summary, the addition of a twist to a fermion tangle or to a photon strand changes their phase, and thus represents a gauge interaction. We have shown that core fluctuations induced by twists produce a U(1) gauge symmetry. Electromagnetic field energy and particle energy are U(1) invariant. In short, the strand model implies that *the gauge group of quantum electrodynamics is U(1)*. With this result, we are now able to deduce the full Lagrangian of QED.

The Lagrangian of QED

Given the U(1) gauge invariance of observables, the Lagrangian of quantum electrodynamics, or QED, follows directly, because U(1) gauge invariance is equivalent to minimal coupling. We start from the Lagrangian density $\mathcal{L}$ of a *neutral*, *free*, and relativistic fermion in an electromagnetic field. It is given by

$$\mathcal{L} = \overline{\Psi}(i\hbar c\not{\partial} - c^2 m)\Psi - \frac{1}{4\mu_0} F_{\mu\nu}F^{\mu\nu} \; . \qquad (151)$$

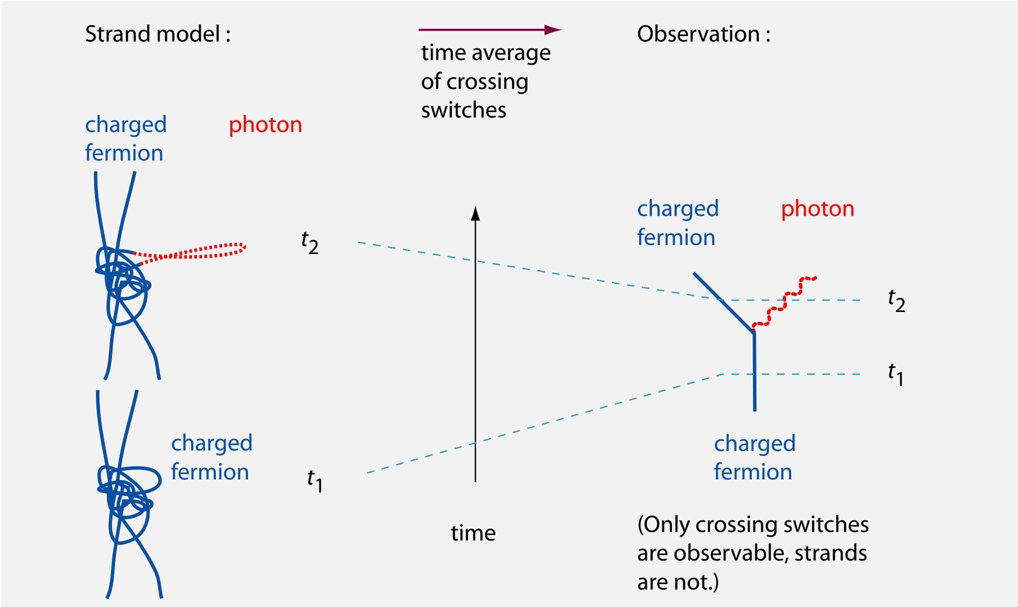


FIGURE 57 The fundamental Feynman diagram of QED and its tangle version.

Page 223 We deduced the fermion term in the chapter of quantum theory, and we deduced the electromagnetic term just now, from the properties of twisted loops.

As we have seen, the strand model implies minimal coupling. This changes the Lagrangian density for a *charged*, i.e., *interacting*, relativistic fermion in the electromagnetic field, into the Lagrangian density of QED:

$$\mathcal{L}_{\mathrm{QED}} = \overline{\Psi}(i\hbar c \not{D} - c^2 m)\Psi - \frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} \, . \tag{152}$$

Here, $\not{D} = \gamma^\sigma \mathrm{D}_\sigma$ is the *gauge covariant derivative* that is defined through minimal coupling to the charge q:

$$\mathrm{D}_\sigma = \partial_\sigma - iqA_\sigma \, . \tag{153}$$

Minimal coupling implies that the Lagrangian density of QED is invariant under U(1) gauge transformations. We will discuss the details of the charge q later on. Page 382

We have thus recovered the Lagrangian density of quantum electrodynamics from strands. Strands thus reproduce the most precisely tested theory of physics.

Feynman diagrams and renormalization

Feynman diagrams are abbreviations of formulas to calculate effects of quantum electrodynamics in perturbation expansion. Feynman diagrams follow from the Lagrangian of QED. All Feynman diagrams of QED can be constructed from one fundamental diagram, shown on the right-hand side of Figure 57. Important Feynman diagrams are shown on the left-hand sides of Figure 58 and of Figure 59.

In the strand model, the fundamental Feynman diagram can be visualized directly

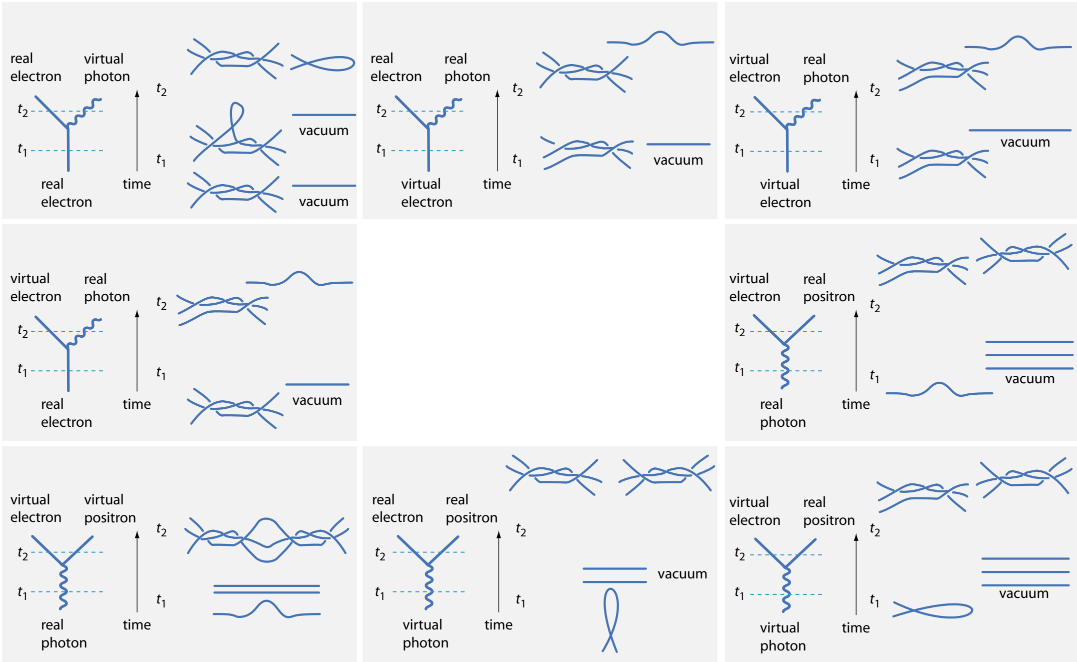


FIGURE 58 The different variations of the fundamental Feynman diagram of QED and their tangle versions.

in terms of strands, as shown on the left-hand side of Figure 57. This is the same diagram that we have explored right at the start of the section on electrodynamics, when we defined electrodynamics as twist exchange. (The precise tangles for the charged fermions will be deduced later on.) Since all possible Feynman diagrams are constructed from the fundamental diagram, the strand model allows us to interpret all possible Feynman diagrams as strand diagrams. For example, the strand model implies that the vacuum is full of virtual particle-antiparticle pairs, as shown in Figure 59.

Page 229
Page 314

In quantum field theory, Lagrangians must not only be Lorentz and gauge invariant, but must also be renormalizable. The strand model makes several statements on this issue. At this point, we focus on QED only; the other gauge interactions will be treated below. The strand model reproduces the QED Lagrangian, which is renormalizable. Renormalizability is a natural consequence of the strand model in the limit that strand diameters are negligible. The reason for renormalizability that the strand model reproduces the single, fundamental Feynman diagram of QED, without allowing other types of diagrams.

The twist deformations underlying the strand model for QED also suggest new ways to calculate higher order Feynman diagrams. Such ways are useful in calculations of g-factors of charged particles, as shown in the next section. In particular, the strand model for QED, as shown in Figure 57, implies that higher order QED diagrams are simple *strand deformations* of lower order diagrams. Taking statistical averages of strand deformations up to a given number of crossings thus allows us to calculate QED effects up to a given order in the coupling. The strand model thus suggests that non-perturbative calculations are possible in QED.

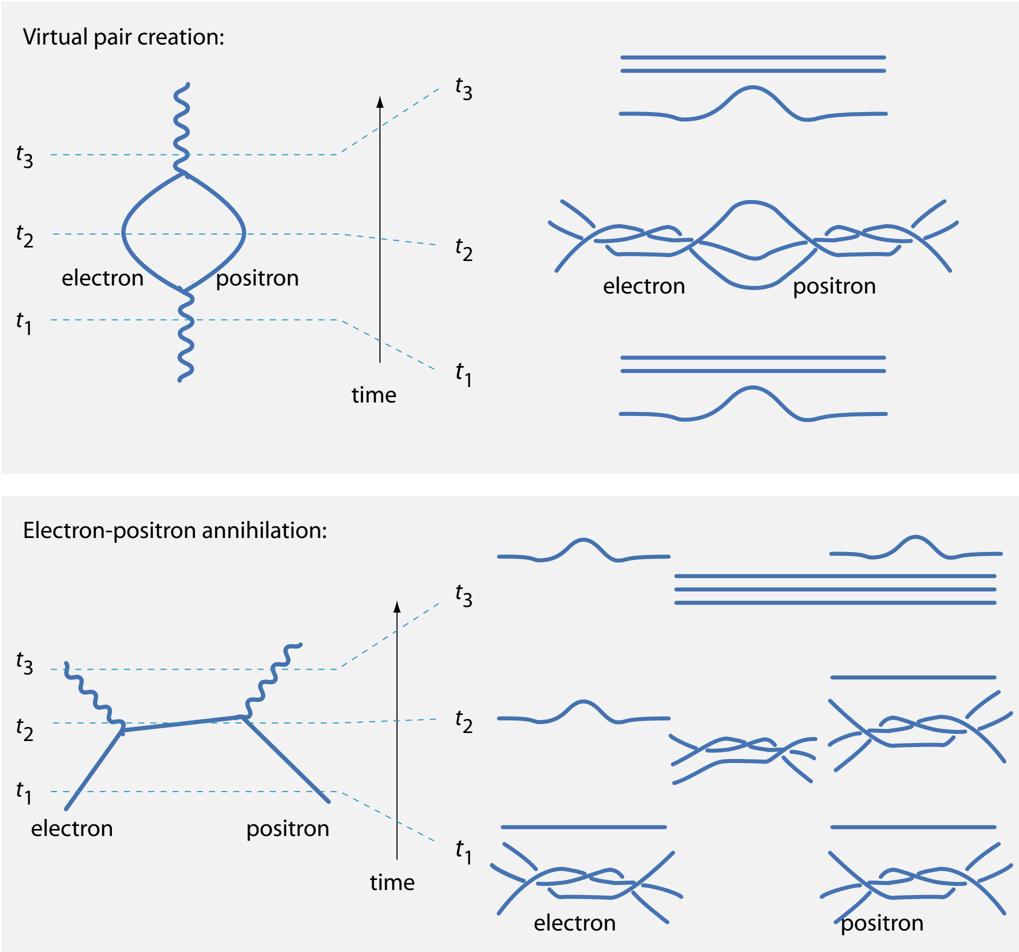


FIGURE 59 Some Feynman diagrams of QED with their tangle versions.

For precise non-perturbative calculations, the effective diameter of the strands must be taken into account. The diameter eliminates the Landau pole and all ultraviolet divergences of QED. In the strand model, the vacuum energy of the electromagnetic field is automatically zero. In other words, the strand model eliminates all problems of QED; in fact, QED appears as an approximation of the strand model for negligible strand diameter. In passing, we thus predict that perturbation theory for QED is valid and *converges* if the strand model, and in particular the finite strand diameter, is taken into account. (The diameter is the only gravitational influence predicted to affect QED.) However, we do not pursue these topics in the present text.

The strand model also suggests that the difference between renormalized and unrenormalized mass and charge is related to the difference between minimal and non-minimal crossing switch number, or equivalently, between tangle deformations with few and with many crossings, where strands are deformed on smaller distance scales. In other terms, unrenormalized quantities – the so-called *bare* quantities at Planck energy – can be imagined as those deduced when the tangles are pulled tight, i.e., pulled to Planck distances, whereas renormalized mass and charge values are those deduced for particles surrounded by many large-size fluctuations.

The strand model also suggests a visualization for the cut-off used in QED. The cut-off is a characteristic energy or length used in intermediate calculations. In the strand model, the cut-off corresponds to the size of the image.

In summary, the strand model provides a new underlying picture or mechanism for Feynman diagrams. The strand model does not change any physical result at any experimentally accessible energy scale. In particular, the measured change or 'running' with Page 386 energy of the fine structure constant and of the masses of charged particles are reproduced by the strand model, because Feynman diagrams of all orders are reproduced up to energies just below the Planck scale. Deviations between QED and the strand model are only expected near the Planck energy, when tangles of Planck diameter are pulled tight.

The anomalous magnetic moment

The anomalous magnetic moment g of the electron and of the muon is given by the well-known expression

$$\frac{g}{2} = 1 + \frac{\alpha}{2\pi} - O(\alpha^2) \,, \tag{154}$$

where $g/2$ is half the so-called g-factor, with a measured value of 1.00116(1), and α is the fine structure constant, with a measured value of 1/137.036(1). Julian Schwinger discovered this expression in 1948; the involved calculations that led Schwinger to this and similar results in quantum field theory earned him the 1965 Nobel Prize in Physics. The result is also inscribed on the memorial marker near his grave in Mount Auburn Cemetery. The strand model proposes an intuitive explanation for this result.

Generally speaking, the factor $g/2$ describes the ratio between the 'mechanical' or 'geometric' rotation frequency – the rotation of the particle *mass* that leads to spin – and 'magnetic' rotation frequency – the rotation of the particle *charge* that leads to the magnetic moment. More precisely, the definition of the g-factor of a particle with charge e and mass m is

$$\frac{g}{2} = \frac{\mu/e}{S/m} \,. \tag{155}$$

Here, μ is the magnetic moment and S is the intrinsic angular momentum, or spin.

The *mechanical* or *geometric* rotation frequency is related to the ratio of the intrinsic angular momentum L and the mass m. Using the definitions from classical physics, we have $\boldsymbol{S}/m = \boldsymbol{r} \times \boldsymbol{v}$. The *magnetic* rotation frequency is related to the ratio of the magnetic moment $\boldsymbol{\mu}$ and the electric charge e. Classically, this ratio is $\boldsymbol{\mu}/e = \boldsymbol{r} \times \boldsymbol{v}$. Therefore, in classical physics – and also in the first order of the Pauli–Dirac description of the electron – the two rotation frequencies coincide, and the factor $g/2$ is thus equal to 1. However, as mentioned, both experiment and QED show a slight deviation of $g/2$ from unity, called the *anomalous* magnetic moment.

In the strand model, the geometric or mechanical rotation of a charged elementary particle is due to the rotation of the tangle core as a rigid whole, whereas the magnetic Page 227 rotation also includes phase changes due to the *deformations of the tangle core*. In par-

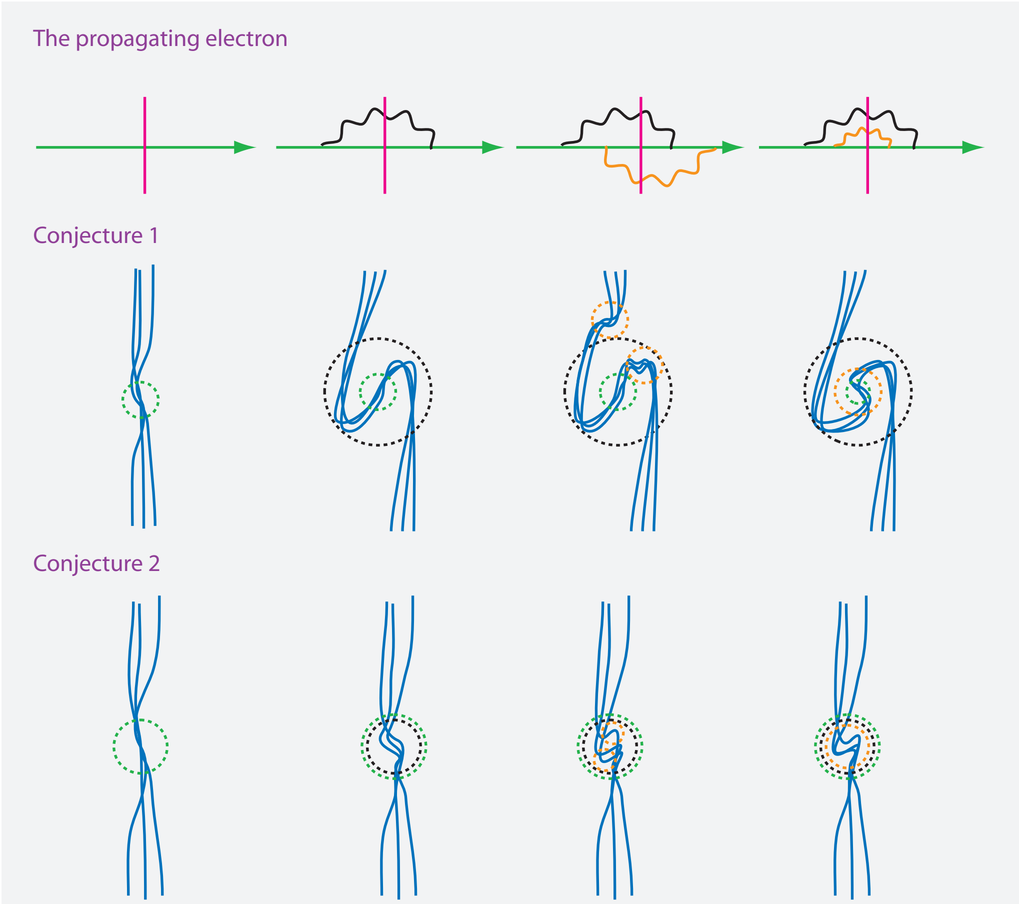


FIGURE 60 Two conjectured correspondences between the Feynman diagrams of quantum electrodynamics and the strand model for a propagating free electron. The lower strand model configurations are shown for a single instant – marked in magenta – of the electron propagator drawn above them. (For simplicity, the external field is not drawn.) In the first conjecture, the loops of the belt trick are conjectured to correspond to the virtual photons in the propagator and to be responsible for the anomalous magnetic moment. In the second conjecture, the deformations of the core correspond to the virtual photons.

ticular, the magnetic rotation of a charged elementary particle includes phase changes due to emission and reabsorption of virtual photons, i.e., of twisted loops.

In nature, the probability of the emission and reabsorption of a photon is determined by the fine structure constant α. The emission and reabsorption process leads to an additional angle that makes the 'magnetic' rotation angle differ from the 'mechanical' rotation angle. Since the fine structure constant describes the rotation of the phase due to virtual photon exchange, the emission and reabsorption of a virtual photon leads to an angle difference, and this angle difference is given by the fine structure constant itself. The ratio between the purely mechanical or geometric and the full magnetic rotation frequency is therefore not one, but increased by the ratio between the additional angle α and 2π. This is Schwinger's formula.

In short, the strand model reproduces Schwinger's celebrated formula for the anomalous magnetic moment almost from thin air. The strand model also implies that Schwinger's formula is valid for *all* charged elementary particles, independently of their mass; this is indeed observed. Higher order corrections also appear naturally in the strand model. Finally, the strand model implies that the complete expression, with all orders included, *converges*, because the full result is due to the shape and dynamics of the tangle core. The discussions about the existence of the perturbation limit in QED are thus laid to rest.

If we look into the details, it might be that the belt trick itself is at the origin of the anomalous magnetic moment. A conjecture for this connection is proposed and illustrated in Figure 60: if the two loops formed by the belt trick are seen as virtual photons, Challenge 156 e the factor $2\alpha/4\pi$ arises naturally. So do the higher-order terms. This explanation would relate the belt trick directly to the additional magnetic rotation angle. However, it might also be that this correspondence of the strand images in the figure to the upper diagrams is not fully correct. The topic is subject of research.

A second conjecture is also given in Figure 60. The virtual photons could correspond to deformations of the tangle core. This conjecture is more in line with the distinction Page 227 between gravity and gauge interactions given above, where it was stated that gravity is due to tail deformations and gauge interactions are due to core deformations. This conjecture is more in line with the distinction between a geometric and a magnetic rotation: the geometric rotation would be due to the rigid rotation of the tangle core, and the magnetic rotation would be due to an additional effect due to core deformation.

Both conjectures on the origin of the g-factor imply that $1 < g/2 < 2$; in fact, we can even argue, using $\alpha < 1$, that the strand model implies

$$1 < g/2 < 1 + \frac{1}{2\pi} . \qquad (156)$$

This is not a new result; it is already implied by ordinary quantum field theory. However, the strand description of particle rotation suggests a way to calculate the g-factor and Page 393 the fine structure constant. We will explore this below.

Maxwell's equations

Ref. 183 The strand model of charge and photons reproduces Maxwell's equations . But strands also allow us to visualize and check Maxwell's field equations of classical electrodynamics directly. The equations are:

$$\begin{aligned}
\nabla\, \mathbf{E} &= \frac{\rho}{\varepsilon_0} , \\
\nabla\, \mathbf{B} &= 0 , \\
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} , \\
\nabla \times \mathbf{B} &= \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J} . \qquad (157)
\end{aligned}$$

▪ The first of these equations is satisfied whatever the precise mechanism at the basis of twisted loop emission by electric charges may be. Indeed, any mechanism in which a charge randomly sends out or swallows a twisted handle yields a $1/r^2$ dependence for the electrostatic field and the required connection between charge and the divergence of the electric field. This is not a deep result: any spherically-symmetric system that randomly emits or swallows some entity produces the equation, including the underlying inverse-square dependence. The result can also be confirmed in another, well-known way. In any exchange interaction between two charges, the exchange time is proportional to their distance apart r; in addition, quantum theory states that the exchanged momentum is inversely proportional to the distance r. Therefore, the force, or momentum per unit time, varies as $1/r^2$. This relation is valid independently of the underlying motion of the twisted loops, because space has three dimensions: all localized sources automatically fulfil the inverse square dependence.

The constant on the right-hand side of the first equation results from the definition of the units; in the language of the strand model, the constant fixes the twisted loop emission rate for an elementary charge.

▪ The second of the field equations (157) expresses the lack of magnetic charges. This equation is automatically fulfilled by the strand model, as the definition of the magnetic field with strands does not admit any magnetic sources. In fact, strands suggest that no localized entity can have a magnetic charge. Also this equation is valid independently of the details of the motion of the strands. Again, this is a topological effect.

▪ The third field equation relates the temporal change of the magnetic field to the curl of the electric field. In the strand model, this is satisfied naturally, because a curl in the electric field implies, by construction, a change of the magnetic field, as shown by Figure 53. Again, this relation is valid independently of the details of the motion of the strands, as long as the averaging scale is taken to be large enough to allow the definition of electric and the magnetic fields.

Page 234

▪ The most interesting equation is the last of the four Maxwell equations (157): in particular, the second term on the right-hand side, the dependence on the charge current. In the description of electrodynamics, the charge current **J** appears with a positive sign and with no numerical factor. (This is in contrast to linearized gravity, where the current has a numerical factor and a negative sign.) The positive sign means that a larger current produces a larger magnetic field. The strand model reproduces this factor: strands lead to an effect that is proportional both to charge (because more elementary charges produce more crossing flows) and to speed of movement of charge (large charge speed lead to larger flows). Because of this result, the classical photon spin, which is defined as L/ω, and which determines the numerical factor, namely 1, that appears before the charge current **J**, is recovered. Also this connection is obviously independent of the precise motion of the underlying strands.

The first term on the right-hand side of the fourth equation, representing the connection between a changing electric field and the curl of the magnetic field, is automatically in agreement with the model. This can again be checked from Figure 53 – and again, this is a topological effect, valid for any underlying strand fluctuation. As an example, when a capacitor is charged, a compass needle between the plates is deflected. In the strand model, the accumulating charges on the plates lead to a magnetic field. The last of Maxwell's equations is thus also confirmed by the strand model.

Page 234

In summary, the strand model reproduces Maxwell's equations. However, this is not a great feat. Maxwell-like equations appear in many places in field theory, for example in solid-state physics and hydrodynamics. Mathematical physicists are so used to the appearance of Maxwell-like equations in other domains that they seldom pay it much attention. The real tests for any model of electrodynamics, quantum or classical, are the deviations that the model predicts from electrodynamics, especially at high energies.

Curiosities and fun challenges about QED

Challenge 157 e Can you show that the calculation of the vacuum energy density of an infinite flat vacuum, when using strands, yields exactly zero, as expected?

* *

Challenge 158 e Can you confirm that the strand model of quantum electrodynamics does not violate charge conjugation C nor parity P at any energy?

* *

Challenge 159 e Can you confirm that the strand model of quantum electrodynamics conserves colour and weak charge at all energies, using the results of the next sections?

* *

Challenge 160 e Can you determine whether the U(1) gauge group deduced here is that of electrodynamics or that of weak hypercharge?

* *

Challenge 161 d Can you find a measurable deviation of the strand model from QED?

Summary on QED and experimental predictions

In the strand model, photons are single, helically twisted strands, randomly exchanged between charges; charges are chiral tangles, and therefore they effectively emit and absorb real and virtual photons. This is the complete description of QED using strands.

In particular, we have shown that Reidemeister I moves – or twists – of tangle cores lead to U(1) gauge invariance, Coulomb's inverse square relation, Maxwell's equations of electrodynamics and to Feynman diagrams. In short, we have deduced all experimental properties of quantum electrodynamics, except one: the strength of the coupling. Despite this open point, we have settled one line of the millennium list of open issues: we know the origin of the electromagnetic interaction and of its properties. Page 164

Is there a difference between the strand model and quantum electrodynamics? The precise answer is: there are *no measurable* differences between the strand model and QED. For example, the g-factor of the electron or the muon predicted by QED is not changed by the strand model. The U(1) gauge symmetry and the whole of QED remain valid at all energies. There are no magnetic charges. There are no other gauge groups. QED remains exact in all cases – as long as gravity plays no role.

The strand model prediction of a lack of larger gauge symmetries is disconcerting. There is thus *no* grand unification in nature; there is no general gauge group in nature, be it SU(5), SO(10), E6, E7, E8, SO(32) or any other. This result indirectly also rules out

Ref. 184 supersymmetry and supergravity. This unpopular result contrasts with many cherished habits of thought.

In the strand model, the equivalence of Feynman diagrams and strand diagrams implies that deviations of the strand model from QED are expected *only* when gravity starts to play a role. The strand model predicts that this will only happen just near the Planck energy $\sqrt{\hbar c^5/4G}$. At lower energies, QED is predicted to remain valid.

The strand model also confirms that the combination of gravity and quantum theory turns all Planck units into *limit* values, because there is a maximum density of strand crossings in nature, due to the fundamental principle. In particular, the strand model confirms the maximum electric field value $E_{\text{max}} = c^4/4Ge \approx 1.9 \cdot 10^{62}$ V/m and a maximum magnetic field value $B_{\text{max}} = c^3/4Ge \approx 6.3 \cdot 10^{53}$ T. So far, these predictions are not in contrast with observations.

Thus the strand model predicts that approaching the electric or magnetic field limit values – given by quantum gravity – is the only option to observe deviations from QED. But measurements are not possible in those domains. Therefore we can state that there are no measurable differences between the strand model and QED.

Our exploration of QED has left open only two points: the calculation of the electromagnetic coupling constant and the determination of the spectrum of possible tangles for the elementary particles. Before we clarify these points, we look at the next Reidemeister move.

THE WEAK NUCLEAR INTERACTION AND THE SECOND REIDEMEISTER MOVE

In nature, the weak interaction is the result of the absorption and the emission of massive spin-1 bosons that form a broken weak triplet. The W and the Z bosons are emitted or absorbed by particles with weak charge; these are the left-handed fermions and right-handed antifermions. In other words, the weak interaction breaks parity P maximally. The W boson has unit electric charge, the Z boson has vanishing electric charge. The emission or absorption of W bosons changes the particle type of the involved fermion. The weak bosons also interact among themselves. All weakly charged particles are massive and move slower than light. The Lagrangian of matter coupled to the weak Page 255 field has a broken SU(2) gauge symmetry. There are fundamental Feynman diagrams with triple and with quartic vertices. The weak coupling constant is determined by the electromagnetic coupling constant and the weak boson masses; its energy dependence is fixed by renormalization. The Higgs boson ensures full consistency of the quantum field theory of the weak interaction.

The previous paragraph summarizes the main observations about the weak interaction. More precisely, all observations related to the weak interaction are described by its Lagrangian. Therefore, we need to check whether the weak interaction Lagrangian follows from the strand model.

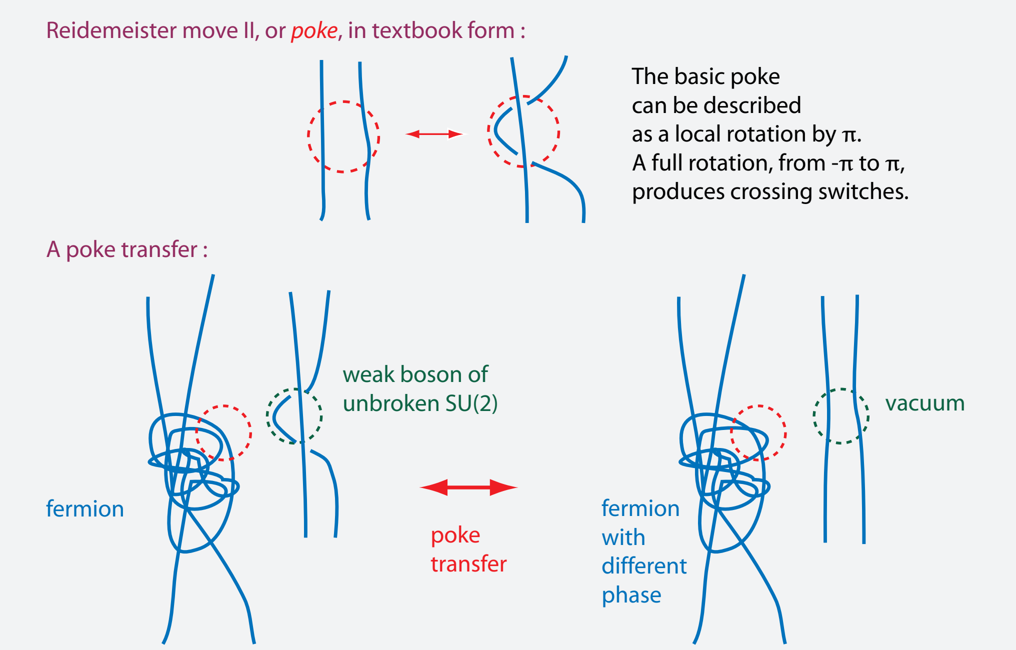


FIGURE 61 *Poke transfer* is the basis of the weak interaction in the strand model. No strand is cut or reglued; the transfer occurs only through the excluded volume due to the impenetrability of strands.

Strands, pokes and SU(2)

Page 225 As explained above, any gauge interaction involving a fermion is a deformation of the tangle core that changes the phase and rotation of the fermion tangle. We start directly with the main definition.

▷ The ***weak interaction*** is the transfer of a poke, i.e., the transfer of a Reidemeister II move, between two particles. An illustration is given in Figure 61. Strands are not cut in this process; they simply transfer the deformation as a result of their impenetrability.

Strands describe the weak interaction as exchange of pokes. In tangle cores, the *basic* pokes induce local rotations by an angle π, as shown in Figure 62: each basic poke rotates the region enclosed by the dotted circle. A full poke produces two crossings. There are *three*, linearly independent, basic pokes, in three mutually orthogonal directions. The three basic pokes τ_x, τ_y and τ_z act on the local region in the same way as the three possible mutually orthogonal rotations act on a belt buckle. For completeness, we note that the following arguments do not depend on whether the two strands involved in a poke are parallel, orthogonal, or at a general angle. The following arguments also do not depend on whether the pokes are represented by deforming *two* strands or only *one* strand. Both cases lead to crossing switches, for each possible poke type. (Challenge 162 e)

Figure 62 illustrates that the product of two different basic pokes gives the third basic poke, together with a sign – which depends on whether the sequence is cyclic or not – (Challenge 163 e)

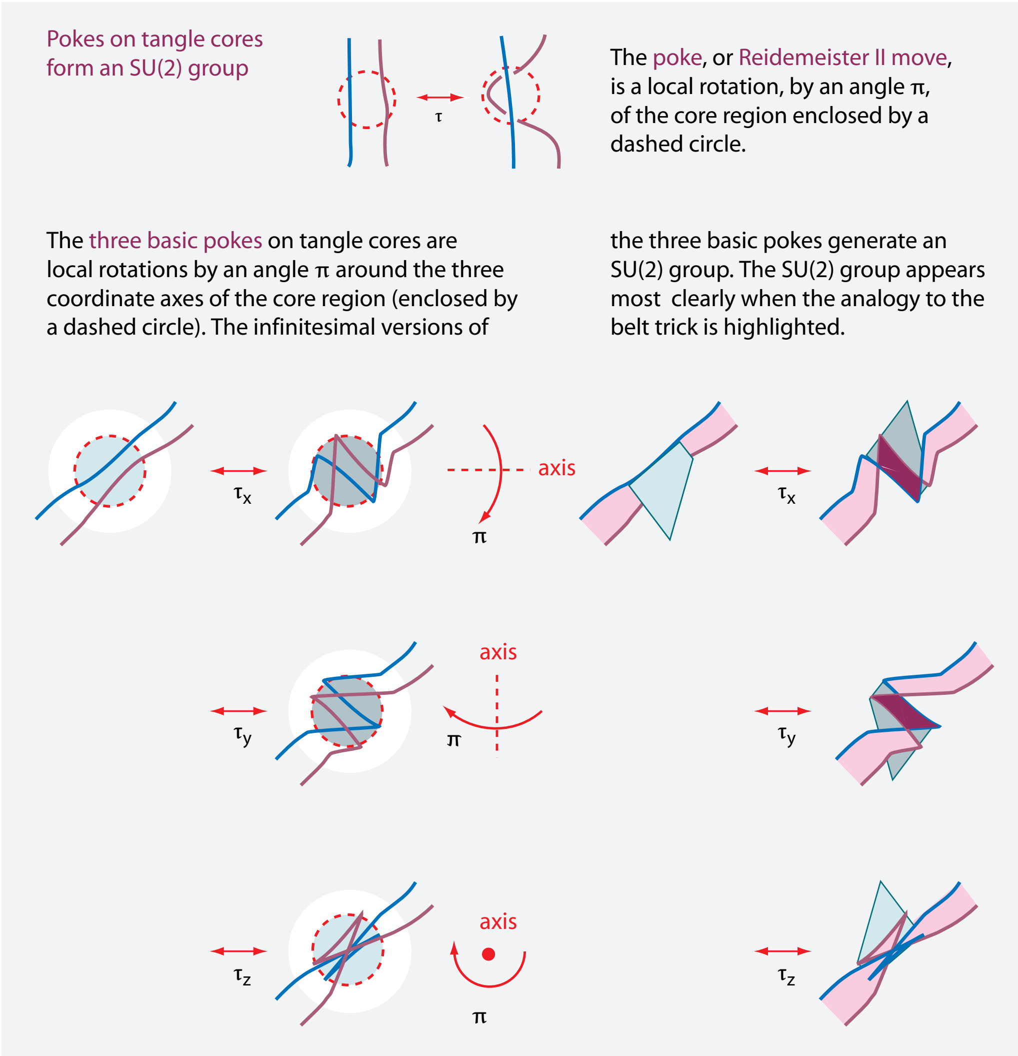


FIGURE 62 How the set of all pokes – the set of all deformations induced on tangle cores by the weak interaction – forms an SU(2) gauge group: the three pokes lead to the belt trick, illustrated here with a pointed buckle and two belts. The illustrated deformations of two strands represent the three unbroken weak vector bosons.

and a factor of i. Using the definition of -1 as a local rotation of the buckle region by Page 191 2π, we also find that the square of each basic poke is -1. In detail, we can read off the following multiplication table for the three basic pokes:

$$\begin{array}{c|ccc} \cdot & \tau_x & \tau_y & \tau_z \\ \hline \tau_x & -1 & i\tau_z & -i\tau_y \\ \tau_y & -i\tau_z & -1 & i\tau_x \\ \tau_z & i\tau_y & -i\tau_x & -1 \end{array} \tag{158}$$

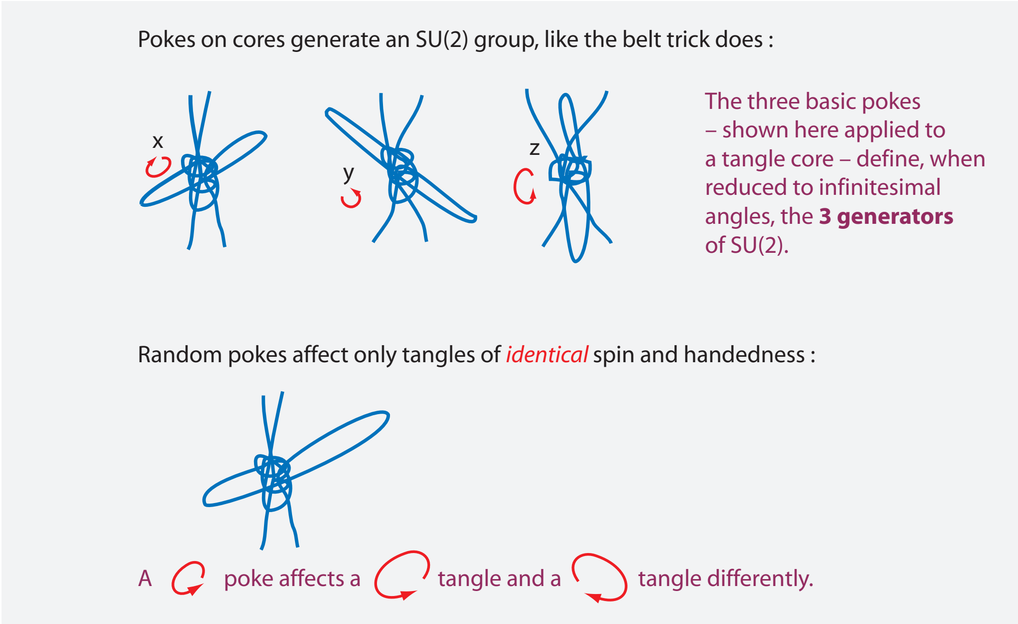


FIGURE 63 The three basic pokes and weak charge in the strand model.

In other terms, the three basic pokes – and in particular also their infinitesimal versions – behave like the generators of an SU(2) group. Because pokes can be seen as local rotations of a buckle region, they can be generalized to arbitrary angles. Such arbitrary pokes can be concatenated. We thus find that arbitrary pokes form a full SU(2) group. This is the reason for their equivalence with the belt trick.

The different gauge choices for a particle are not illustrated in Figure 62. The gauge choices arise from the different ways in which the basic pokes τ_x, τ_y and τ_z can be assigned to the set of deformations that describe the belt trick.

In summary, we can state that in any definition of the phase of a tangled fermion core, there is an SU(2) gauge freedom; in addition, there exists an interaction with SU(2) gauge symmetry. In other words, the strand model implies, through the second Reidemeister move, *the existence of the unbroken weak interaction with a gauge group SU(2).*

Weak charge and parity violation

A particle has weak charge if, when subject to many random pokes, a non-zero average phase change occurs. Surrounded by a bath of strands that continuously induce random pokes, not all tangles will change their phase on a long-time average: only tangles that lack symmetry will. One symmetry that must be lacking is spherical symmetry. Therefore, only tangles whose cores lack *spherical symmetry* have the chance to be influenced by random pokes. Since all tangles, independently of their core details, lack spherical symmetry, all such tangles, i.e., all massive particles, are candidates to be influenced, and thus are candidates for weakly charged particles. We therefore explore them in detail now.

If a tangle is made of *two or more* linked strands, it represents a massive spin-1/2

Page 298 particle (except for a simple twist, which represents the graviton). All such fermion cores lack spherical and cylindrical symmetry. When a fermion spins, two things happen: the core rotates and the belt trick occurs, which untangles the tails. Compared to the Page 220 direction of motion, the rotation and the untangling can be either left-handed or right-handed.

Every poke is a shape transformation of the core with a preferred handedness. The chirality is of importance in the following.

A particle has weak charge if random pokes lead to a long-time phase change. In order to feel any average effect when large numbers of random pokes are applied, a core must undergo different effects for a poke and its reverse. As already mentioned, this requires a lack of core symmetry. Whenever the core has no symmetry, non-compensating phase effects will occur: if the core rotation with its tail untangling and the poke are of the same handedness, the phase will increase, whereas for opposite handedness, the phase will decrease a bit less.

▷ Non-vanishing ***weak charge*** for fermions appears only for tangle cores whose handedness leads to average poke effects.

In other words, the strand model predicts that random pokes will only affect a core if the core handedness and the randomly applied belt trick are of the *same* handedness. In physical terms, random pokes will only affect left-handed particles or right-handed antiparticles. Thus, the strand model predicts that *the weak interaction violates parity maximally*, This is exactly as observed. In other terms, weak charge and the parity violation of the weak interaction are consequences of the belt trick. This relation is summarized in Figure 63.

If an elementary particle is described by a *two tangled* strands, we expect it to be influenced by average pokes. Such tangle cores are spin-1 bosons; their cores lack spherical and cylindrical symmetry. The core rotation will induce a left-right asymmetry that will lead to a higher effect of a poke than of its reverse. Two-stranded particles are thus predicted to carry weak charge. We therefore expect that quarks – to be explored below – and the weak bosons themselves interact weakly.

Because the weak bosons interact weakly, the strand model implies that the weak interaction is a *non-Abelian* gauge theory, as is observed.*

If a tangle is made of a *single unknotted* strand, it is not affected by random pokes. The strand model thus predicts that the photon has no weak charge, as is observed. The same also holds for gluons.

The strand definition of weak charge leads to two conclusions that can be checked by experiment. First, all electrically charged particles – having cores that are chiral and thus lack cylindrical symmetry – are predicted to be weakly charged. Secondly, in the strand model, only massive particles interact weakly; in fact, *all* massive particles interact weakly, because their cores lack cylindrical symmetry. In other words, all weakly

* Non-Abelian gauge theory was introduced by Wolfgang Pauli. In the 1950s, he explained the theory in series of talks. Two physicists, Yang Chen Ning and Robert Mills, then wrote down his ideas. Yang later received the Nobel Prize in Physics with Lee Tsung Dao for a different topic, namely for the violation of parity of the weak interaction.

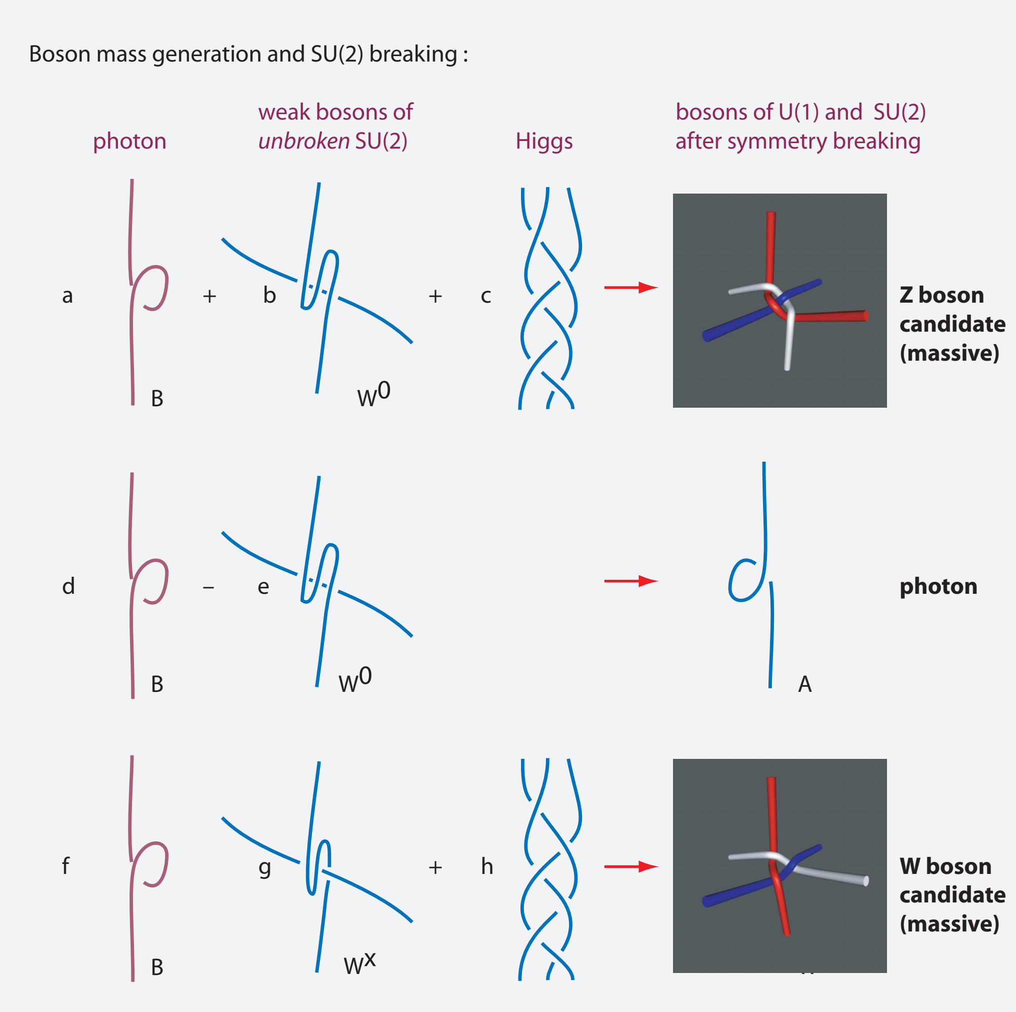


FIGURE 64 Poke-inducing strand motions (left) become massive weak vector tangles (right) through symmetry breaking and tail braiding. Tail braiding is related to the Higgs boson, whose tangle model will be clarified later on.

charged particles move more slowly than light and vice versa. Both conclusions agree with observation.

In summary, all properties of weak charge found in nature are reproduced by the tangle model.

Weak bosons

Gauge bosons are those particles that are exchanged between interacting fermions: gauge bosons induce phase changes of fermions. This implies that the (unbroken) weak bosons are the particles* that induce the three poke moves:

* This reworked strand model of the W and Z bosons arose in 2015.

▷ ***Weak, unbroken, intermediate bosons*** are described by double strands. Illustrations are given in and Figure 62 and Figure 64.

Single strands that induce phase changes in fermions interacting weakly are shown on the left side of Figure 64. They correspond to the three basic pokes τ_x, τ_y and τ_z.

We note two additional points. First of all, the (unbroken) bosons could also be described by the motion of a single strand in a strand group. This makes them *spin 1* particles.

Furthermore, unknotted tangles are *massless*. In the strand model, tangles that induce pokes *differ* from the massive weak intermediate bosons, shown on the right of Figure 64. This difference is due to the *breaking* of the SU(2) gauge symmetry, as we will find out soon.

The Lagrangian of the unbroken SU(2) gauge interaction

The energy of the weak field is given by the density of weak gauge boson strands. As long as the SU(2) symmetry is not broken, the energy of the weak field and the energy of fermions are both SU(2) invariant. As a consequence, we are now able to deduce a large part of the Lagrangian of the weak interaction, namely the Lagrangian for the case that the SU(2) symmetry is unbroken.

As long as SU(2) is unbroken, the vector bosons are described as unknotted tangles that induce pokes, as shown on the left of Figure 64. There are three such bosons. Since they can be described by a single strand that moves, they have spin 1; since they are unknotted, they have zero mass and electric charge.

Energy is the number of crossing switches per time. As long as SU(2) is unbroken and the weak bosons are massless, the energy of the weak boson field and thus their Page 234 Lagrangian density is given by the same expression as the energy of the photon field. In particular, the strand model implies that energy density is quadratic in the field intensities. We only have to add the energies of all three bosons together to get:

$$\mathcal{L} = -\frac{1}{4}\sum_{a=1}^{3} W^a{}_{\mu\nu} W_a{}^{\mu\nu} \, , \tag{159}$$

This expression is SU(2) gauge invariant. Indeed, SU(2) gauge transformations have no effect on the number of crossing switches due to weak bosons or to the motion of pokes. Thus, gauge transformations leave weak field intensities and thus also the energy of the weak fields invariant, as observed.

We can now write down the Lagrangian for weakly charged fermions interacting with the weak vector bosons. Starting from the idea that tangle core deformations lead to phase redefinitions, we have found that pokes imply that the *unbroken* weak Lagrangian density for matter and radiation fields is SU(2) gauge invariant. In parallel to electrodynamics we thus get the Lagrangian

$$\mathcal{L}_{\text{unbroken weak}} = \sum_f \overline{\Psi}_f (i\hbar c \not{D} - m_f c^2)\Psi_f - \frac{1}{4}\sum_{a=1}^{3} W^a{}_{\mu\nu} W_a{}^{\mu\nu} \, , \tag{160}$$

where $\not{D}$ is now the SU(2) gauge covariant derivative and the first sum is taken over all fermions. In this Lagrangian, only the left-handed fermions and the right-handed antifermions carry weak charge. This Lagrangian, however, does *not* describe nature: the observed SU(2) breaking is missing.

SU(2) breaking

In nature, the weak interaction does *not* have an SU(2) gauge symmetry. The symmetry is only approximate; is said to be *broken*. The main effect of SU(2) symmetry breaking are the non-vanishing – and different – masses for the W and Z bosons, and thus the weakness and the short range of the weak interaction. In addition, the symmetry breaking implies a *mixing* of the weak and the electromagnetic interaction: it yields the so-called *electroweak* interaction. This mixing is often called electroweak 'unification'.

Page 257

The strand model suggests the following description:

▷ ***Mass generation*** for bosons and the related SU(2) ***symmetry breaking*** are due to *tail braiding* at the border of space. Figure 64 illustrates the idea.

In this description, tail braiding* is assumed to occur at a distance outside the domain of observation; in that region – which can be also the border of physical space – tail braiding is *not* forbidden and *can* occur. The probability of tail braiding is low, because the crossings have first to fluctuate to that distance and then fluctuate back. Nevertheless, the process of tail braiding can take place.

Tail braiding appears *only* in the weak interaction. It does not appear in the other two gauge interactions, as the other Reidemeister moves are not affected by processes at the border of space. In the strand model, this is the reason that only SU(2) is broken in nature. In short, SU(2) breaking is a natural consequence of the second Reidemeister move.

Tail braiding transforms the unbraided, and thus massless, poke strands into the braided, and thus massive W and Z strands. Tail braiding leads to particle cores: therefore is a mass-generating process. The precise mass values that it generates will be determined below. The strand model thus confirms that mass generation is related to the breaking of the weak interaction.

Page 358

Tail braiding mixes the W^0 with the 'original' photon. This is shown in Figure 64. The mixing is due to the topological similarities of the strand models of the two particles. The resulting Z boson is achiral, and thus electrically neutral, as observed. We note that the existence of a neutral, massive Z boson implies that elastic neutrino scattering in matter occurs in nature, as was observed for the first time in 1974. Since any electrically charged particle also has weak charge, the existence of a Z boson implies that any two electrically charged particles can interact both by exchange of photons and by exchange of Z bosons. In other words, SU(2) breaking implies electroweak mixing, or, as is it usually called, electroweak 'unification'.

Page 250
Page 257

Tail braiding takes place in several weak interaction processes, as shown in Figure 67. Tail braiding thus can change particle topology, and thus particle type. The strand model

Page 256

* In the original strand model of the weak bosons, from the year 2008, the role of tail braiding was taken by strand overcrossing.

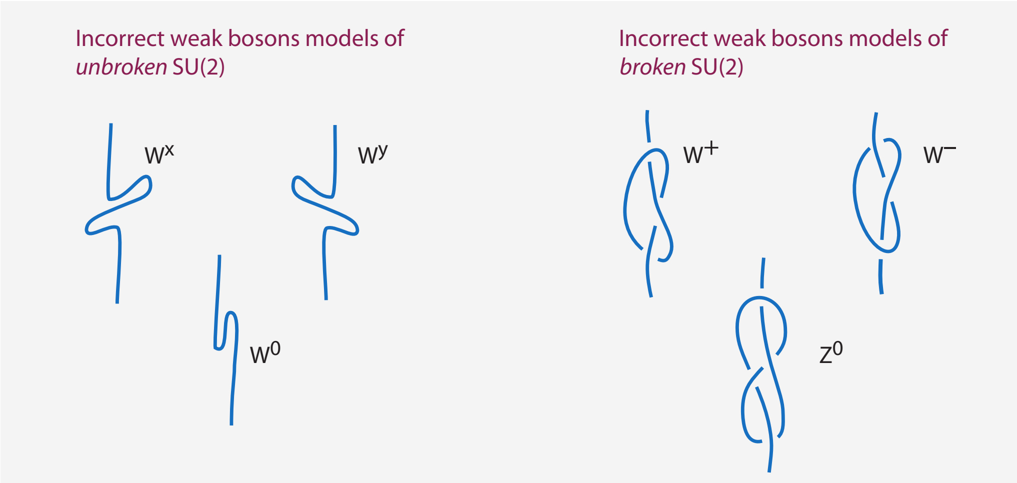


FIGURE 65 The supposed models for the massive weak gauge bosons after symmetry breaking, from 2008 (on the right side), now seen to be incorrect.

thus predicts that the weak interaction *changes* particle flavours (types), as is observed. In fact, the strand model also predicts that *only* the weak interaction has this property. This is also observed.

On the other hand, strands are never cut or glued back together in the strand model, not even in the weak interaction. As a result, the strand model predicts that the weak in-
Page 318 teraction conserves electric charge, spin and, as we will see below, colour charge, baryon number and lepton number. All this is observed.

Tail braiding also implies that the tangles for the Z boson and for the W boson shown above are only the *simplest* tangles associated with each boson; more complicated tangles
Page 358 are higher order propagating states of the same basic open knots. This will be of great importance later on, for the proof that all gauge bosons of nature are already known today.

In summary, the second Reidemeister move leads to *tail braiding*; tail braiding leads to the observed properties of SU(2) symmetry breaking. (Equivalently, the strand model implies that the simplest tangles of the weak interaction bosons show SU(2) symmetry, whereas the more complicated, massive tangles break this symmetry.) The value of the
Page 360 mixing angle and the particle masses have still to be determined. This will be done below.

Open issue: are the W and Z tangles correct?

In 2014, Sergei Fadeev raised an issue: A *tangle* version of the W and Z that does *not* contain any knot and does not require an actual strand overcrossing process at spatial infinity, the strand model would gain in simplicity and elegance. Thinking about the issue, it became clear that such a tangle could occur when vacuum strands were included, as shown above.

In contrast, in 2008, in the first version of the strand model, the W boson after symmetry breaking was thought to be an open overhand knot, and the Z boson an open figure-eight knot.

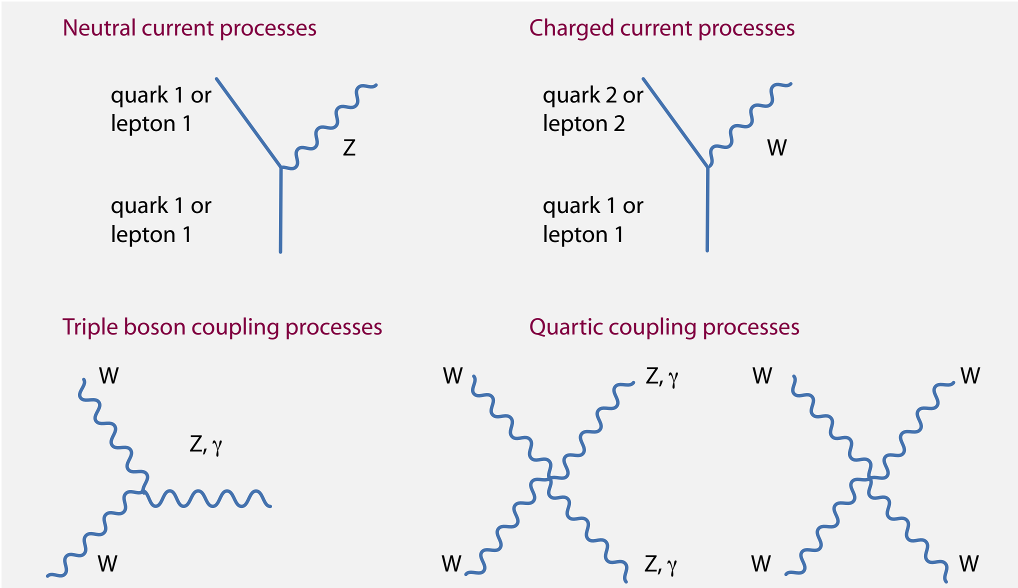


FIGURE 66 The fundamental Feynman diagrams of the weak interaction that do not involve the Higgs boson.

Page 251
Challenge 164 ny

It might well be that the new, 2015/2016 strand models for the two intermediate vector bosons, shown in Figure 64 are still not correct. The possibility remains intriguing and a definitive issue still needs to be found.

The electroweak Lagrangian

We can now use the results on SU(2) symmetry breaking to deduce the *electroweak* Lagrangian density. We have seen that symmetry breaking leaves the photon massless but introduces masses to the weak vector bosons, as shown in Figure 64. The non-vanishing boson masses M_W and M_Z add kinetic terms for the corresponding fields in the Lagrangian.

Due to the symmetry breaking induced by tail braiding, the Z boson results from the mixing with the (unbroken) photon. The strand model predicts that the mixing can be described by an angle, the so-called weak mixing angle θ_w. In particular, the strand model implies that $\cos\theta_w = M_W/M_Z$.

Page 328
Page 376

As soon as symmetry breaking is described by a mixing angle due to tail braiding, we get the known electroweak Lagrangian, though at first without the terms due to the Higgs boson. (We will come back to the Higgs boson later on.) We do not write down the Lagrangian of the weak interaction predicted by the strand model, but the terms are the same as those found in the standard model of elementary particles. There is one important difference: the Lagrangian so derived does not yet contain quark and lepton mixing. Indeed, experiments show that the weak fermion eigenstates are not the same as the strong or electromagnetic eigenstates: quarks mix, and so do neutrinos. The reason for this observation, and the effect that mixing has on the weak Lagrangian, will become clear once we have determined the tangles for each fermion.

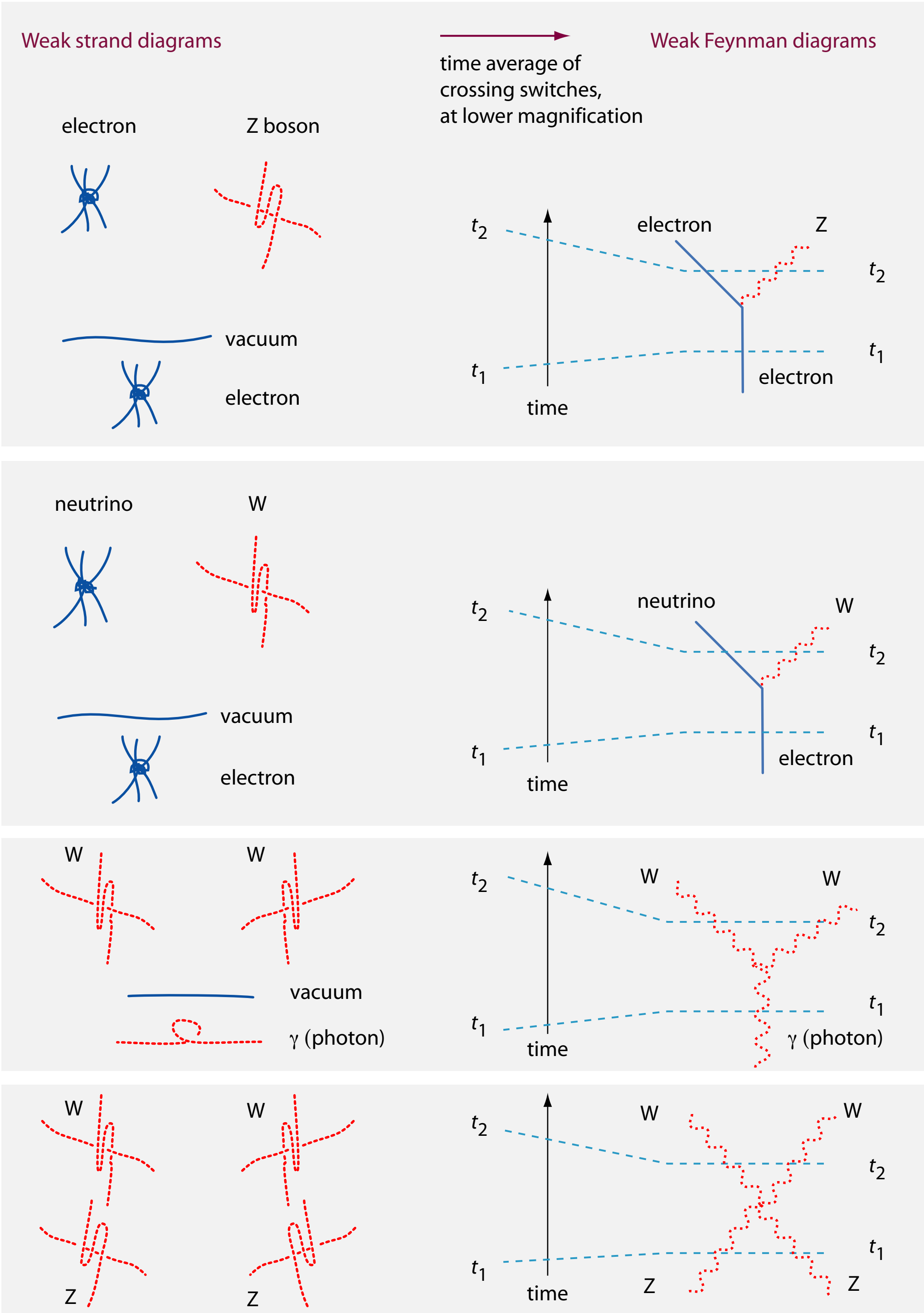


FIGURE 67 The strand model for the fundamental Feynman diagrams of the weak interaction. The tangles for the fermions are introduced later on.

In summary, the strand model implies the largest part of the Lagrangian of the weak interaction. The issue of the Higgs boson is still open, and the electroweak Lagrangian contains a number of constants that are not yet clarified. These unexplained constants are the number of the involved elementary particles, their masses, couplings, mixing angles and CP violation phases, as well as the value of the weak mixing angle.

The weak Feynman diagrams

In nature, the weak interaction is described by a small number of fundamental Feynman diagrams. Those not containing the Higgs boson are shown in Figure 66. These Feynman diagrams encode the corresponding Lagrangian of the weak interaction.

In the strand model, pokes lead naturally to strand versions of the fundamental Feynman diagrams. This happens as shown in Figure 67. We see again that the strand model reproduces the weak interaction: each Feynman diagram is due to a strand diagram for which only crossing switches are considered, and for which Planck size is approximated as zero size. In particular, the strand model does not allow any *other* fundamental diagrams for the weak interaction.

Challenge 165 e

The finite and small number of possible strand diagrams and thus of Feynman diagrams implies that the weak interaction is *renormalizable*. For example, the change or 'running' of the weak coupling with energy is reproduced by the strand model, because the running can be determined through the appropriate Feynman diagrams.

Page 386

Fun challenges and curiosities about the weak interaction

The W boson and its antiparticle are observed to annihilate through the electromagnetic interaction, yielding two or more photons. The tangle model of the weak bosons has a lot of advantages compared to the knot model: The annihilation is much easier to understand.

Challenge 166 e

* *

The strand model, like the standard model of particle physics, predicts that everything about the weak interaction is already known. Nevertheless, the most important weak process, the *decay of the neutron*, is being explored by many precision experiments. The strand model predicts that none of these experiments will yield any surprise.

* *

The strand model makes clear that the weak interaction and the electromagnetic interaction *mix*, but do not unify. There is only electroweak mixing, and *no* electroweak unification, despite claims to the contrary by the Nobel Prize committee and many other physicists. In fact, Sheldon Glashow, who received the Nobel Prize in Physics for this alleged 'unification', agrees with this assessment. So do Richard Feynman and, above all, Martin Veltman, who was also involved in the result; he even makes this very point in his Nobel Prize lecture. The incorrect habit to call electroweak mixing a 'unification' was one of the main reason for the failure of past unification attempts: it directed the attention of researchers in the wrong direction.

Ref. 185

In the strand model, the mixing of the electromagnetic and the weak interaction can be seen as a consequence of knot geometry: the poke generators of the weak interaction also contain twists, i.e., also contain generators of the electromagnetic interaction. In contrast, generators of other Reidemeister moves do not mix among them or with pokes; and indeed, no other type of interaction mixing is observed in nature.

Summary on the weak interaction and experimental predictions

We have deduced the main properties of the weak Lagrangian from the strand model. We have shown that Reidemeister II moves – or pokes – in tangle cores lead to a broken SU(2) gauge group and to massive weak bosons. We found that the deviation from tangle core sphericity plus chirality is weak charge, and that the weak interaction is non-Abelian. We have also shown that the weak interaction naturally breaks parity maximally and mixes with the electromagnetic interaction. In short, we have deduced the main experimental properties of the weak interaction.

Is there a difference between the strand model and the electroweak Lagrangian of the standard model of particle physics? Before we can fully answer the question on deviations between the strand model and the standard model, we must settle the issue of the Higgs boson. This is done later on. Page 328

In any case, the strand model predicts that the broken SU(2) gauge symmetry remains valid at all energies. No other gauge groups appear in nature. The strand model thus predicts again that there is no grand unification, and thus no larger gauge group, be it SU(5), SO(10), E6, E7, E8, SO(32) or any other group. Also this result indirectly rules out supersymmetry and supergravity.

The strand model also predicts that the combination of gravity and quantum theory turns all Planck units into *limit* values, because there is a maximum density of strand crossings in nature, due to the fundamental principle. Therefore, the strand model predicts a *maximum weak field* value given by the Planck force divided by the smallest weak charge. All physical systems – including all astrophysical objects, such as neutron stars, quark starts, gamma-ray bursters or quasars – are predicted to conform to this limit. So far, no observed field value is near this limit, so that the prediction does not contradict observation.

So far, our exploration of the weak interaction has left us with a few open issues: we need to calculate the weak coupling constant and determine the tangle for each particle of the standard model, including the Higgs boson. But we also need to explain weak fermion mixing, CP violation and the masses of all particles. Despite these open points, we have settled another line of the millennium list: we know the origin of the weak interaction and of its main properties. Before we clarify the open issues, we explore the third Reidemeister move. Page 164

The strong nuclear interaction and the third Reidemeister move

In nature, the strong interaction is the result of the absorption and the emission of massless, electrically uncharged, spin-1 gauge bosons that are called *gluons*. Gluons interact with quarks, the only fermions with *colour* charge. Fermions can have three different colour charges, antifermions three different anticolours. Gluons form an octet, are them-

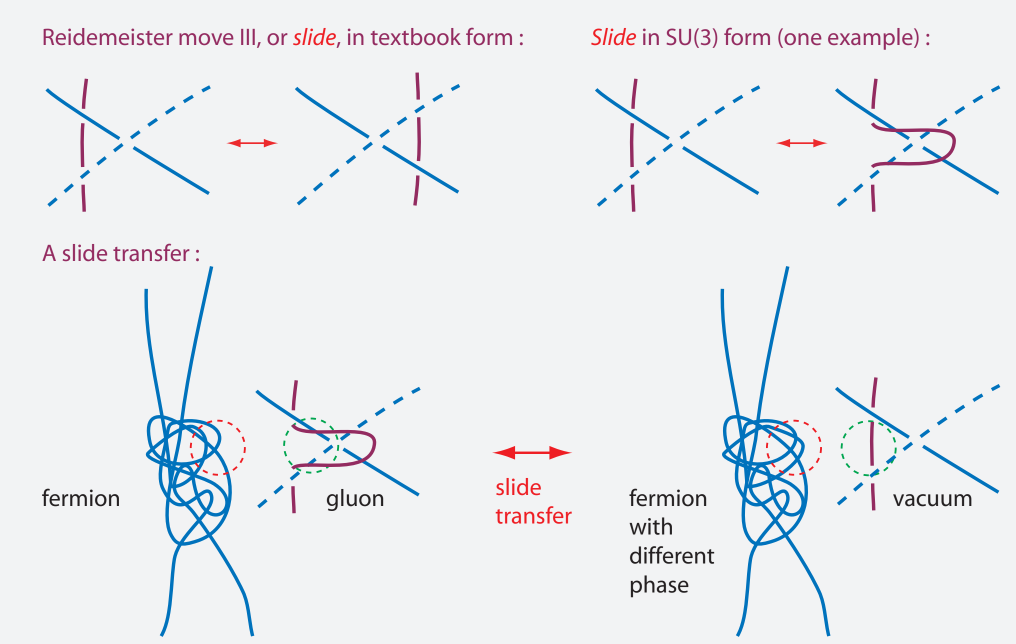


FIGURE 68 A gluon changes the phase of a tangle: *slide transfer* is the basis of the strong interaction in the strand model. During the interaction, no strand is cut or reglued; the transfer occurs purely through the excluded volume that results from the impenetrability of strands. Slide transfer generates a SU(3) gauge group, as explained in the text.

selves colour charged and therefore also interact among themselves. The Lagrangian of quarks coupled to the gluon field has an unbroken SU(3) gauge symmetry. There are three fundamental Feynman diagrams: one for quark-gluon interaction and two for gluon-gluon interactions: a triple and a quartic gluon vertex. The strong coupling constant is about 0.5 at low energy; its energy dependence is determined by renormalization. Its value decreases with increasing energy.

The previous paragraph summarizes the main observations about the strong interaction. All known observations related to the strong interaction, without any known exception, are contained in its Lagrangian. Therefore, we need to show that the strong interaction Lagrangian follows from the strand model.

Strands and the slide, the third Reidemeister move

Page 225 As explained above, interactions of fermions are deformations of the tangle core that change its phase. We start directly by presenting the strand model for the strong interaction.

▷ The ***strong interaction*** is the transfer of *slides*, i.e., the transfer of third Reidemeister moves, between a gluon and a particle. As shown in Figure 68, strands are not cut in this process; gluons simply transfer slide deformations to tangle cores as a result of their impenetrability.

Such a slide transfer will influence the phase of the affected particle tangle. Therefore, slide transfers are indeed a type of interaction.

An introduction to SU(3)

Before we show that slides are responsible for the strong nuclear interaction, we summarize the mathematical properties of the Lie group SU(3). This Lie group is the structure generated by the unitary 3×3 matrices with determinant $+1$. It is a *group*, because matrices can be properly multiplied, because the identity matrix is included, and inverse matrices exist. SU(3) is also a *manifold*; a quick check shows that it has eight dimensions.

Challenge 167 e

In short, SU(3) is a *Lie group*: its elements behave like points on a manifold that can be multiplied. The Lie bracket is the commutator. A general element E of SU(3) can be written as an exponential in the well-known way

$$E = \mathrm{e}^{\sum_{n=1}^{8} \alpha_n i \lambda_n/2} \tag{161}$$

where the eight real parameters α_n can be thought of as the eight coordinates of the group elements on the group manifold. Since SU(3) is compact and simple, these coordinates are best visualized as angles. Of course, i is the imaginary unit. The generators λ_n are complex, traceless and hermitian 3×3 matrices; they are used to define a basis for the group elements. The eight generators are *not* group elements themselves. They describe the structure of the group manifold near the identity matrix; for a Lie group, this local structure defines the full group manifold. Like for any basis, also set of eight generators λ_n is not unique. Of the many possible choices for the generators, the *Gell-Mann matrices* λ_1 to λ_8 are the most commonly used in physics.

The Gell-Mann matrices λ_n, the corresponding group elements D_n for general angles, and the group elements E_n for the finite angle π are given by:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; D_1(\alpha) = \mathrm{e}^{\alpha \mathrm{i} \lambda_1/2} = \begin{pmatrix} \cos\alpha/2 & i\sin\alpha/2 & 0 \\ i\sin\alpha/2 & \cos\alpha/2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$E_1 = \mathrm{e}^{\pi \mathrm{i} \lambda_1/2} = \begin{pmatrix} 0 & i & 0 \\ i & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; D_2(\alpha) = \mathrm{e}^{\alpha \mathrm{i} \lambda_2/2} = \begin{pmatrix} \cos\alpha/2 & \sin\alpha/2 & 0 \\ -\sin\alpha/2 & \cos\alpha/2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$E_2 = \mathrm{e}^{\pi \mathrm{i} \lambda_2/2} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; D_3(\alpha) = \mathrm{e}^{\alpha \mathrm{i} \lambda_3/2} = \begin{pmatrix} \cos\alpha/2 + i\sin\alpha/2 & 0 & 0 \\ 0 & \cos\alpha/2 - i\sin\alpha/2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$E_3 = \mathrm{e}^{\pi \mathrm{i} \lambda_3/2} = \begin{pmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; D_4(\alpha) = \mathrm{e}^{\alpha \mathrm{i} \lambda_4/2} = \begin{pmatrix} \cos\alpha/2 & 0 & i\sin\alpha/2 \\ 0 & 1 & 0 \\ i\sin\alpha/2 & 0 & \cos\alpha/2 \end{pmatrix},$$

$$E_4 = \mathrm{e}^{\pi \mathrm{i} \lambda_4/2} = \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ i & 0 & 0 \end{pmatrix}$$

$$\lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \; D_5(\alpha) = \mathrm{e}^{\alpha \mathrm{i} \lambda_5/2} = \begin{pmatrix} \cos\alpha/2 & 0 & \sin\alpha/2 \\ 0 & 1 & 0 \\ -\sin\alpha/2 & 0 & \cos\alpha/2 \end{pmatrix},$$

$$E_5 = \mathrm{e}^{\pi \mathrm{i} \lambda_5/2} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \; D_6(\alpha) = \mathrm{e}^{\alpha \mathrm{i} \lambda_6/2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha/2 & i\sin\alpha/2 \\ 0 & i\sin\alpha/2 & \cos\alpha/2 \end{pmatrix},$$

$$E_6 = \mathrm{e}^{\pi \mathrm{i} \lambda_6/2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & i \\ 0 & i & 0 \end{pmatrix}$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \; D_7(\alpha) = \mathrm{e}^{\alpha \mathrm{i} \lambda_7/2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha/2 & \sin\alpha/2 \\ 0 & -\sin\alpha/2 & \cos\alpha/2 \end{pmatrix},$$

$$E_7 = \mathrm{e}^{\pi \mathrm{i} \lambda_7/2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

$$\lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$

$$D_8(\alpha) = \mathrm{e}^{\sqrt{3}\alpha \mathrm{i} \lambda_8/2} = \begin{pmatrix} \cos\alpha/2 + i\sin\alpha/2 & 0 & 0 \\ 0 & \cos\alpha/2 + i\sin\alpha/2 & 0 \\ 0 & 0 & \cos\alpha - i\sin\alpha \end{pmatrix},$$

$$E_8 = D_8(\pi) = \begin{pmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -1 \end{pmatrix} . \tag{162}$$

The eight Gell-Mann matrices λ_n are hermitean, traceless and trace-orthogonal. The corresponding group elements D_n and E_n can be thought as the unnormed and normed basis vectors of the group manifold. We note that the definition of E_8 differs from that of the other group elements E_n: it contains an extra factor $\sqrt{3}$. The fourfold concatenation of each matrix $i\lambda_n$ is the identity matrix – except for the case $i\lambda_8$. Instead, the generator λ_8 commutes with λ_1, λ_2 and λ_3 – though not with the other generators.

There is *no* ninth or tenth Gell-Mann matrix. Such a matrix would not be linearly independent from the first eight ones. Indeed, the two matrices deduced from λ_3 using symmetry considerations, namely

$$\lambda_9 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; D_9(\alpha) = \mathrm{e}^{\alpha i \lambda_9/2} = \begin{pmatrix} \cos\alpha/2 - i\sin\alpha/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \cos\alpha/2 + i\sin\alpha/2 \end{pmatrix},$$
$$E_9 = D_9(\pi) = \begin{pmatrix} -i & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & i \end{pmatrix}$$

$$\lambda_{10} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \; D_{10}(\alpha) = \mathrm{e}^{\alpha i \lambda_{10}/2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha/2 + i\sin\alpha/2 & 0 \\ 0 & 0 & \cos\alpha/2 - i\sin\alpha/2 \end{pmatrix},$$
$$E_{10} = D_{10}(\pi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix} \qquad (163)$$

are linear combinations of λ_3 and λ_8; in particular, we have $\lambda_3 + \lambda_9 + \lambda_{10} = 0$ and $\sqrt{3}\lambda_8 + \lambda_9 = \lambda_{10}$. Therefore, λ_9 and λ_{10} are *not* Gell-Mann matrices. (Also two further matrices corresponding to λ_8 in the other two triplets can be defined. The sum of these three matrices is 0 as well.)

The multiplication properties of the Gell-Mann generators λ_1 to λ_8 are listed in Table 10. To make the threefold symmetry more evident, the table also lists the products containing the linearly dependent matrices λ_9 and λ_{10}. Writing the table with the commutators would directly show that the generators form a Lie algebra.

Vol. V, page 226

The *centre* of SU(3) – the subgroup that commutes with all other elements of the group – is Z_3; its threefold symmetry is useful in understanding the behaviour of the group elements and of the generators in more detail.

The group elements E_1 to E_8 listed above share the property that their fourth powers $(E_n)^4$ are the identity matrix. The first matrix triplet E_1, E_2, E_3, the second triplet E_4, E_5, E_9 and the third triplet E_6, E_7, E_{10} each form a SU(2) subgroup. Reflecting the threefold symmetry of its centre, SU(3) contains three linearly independent SU(2) subgroups. The group element E_8 commutes with the first triplet E_1, E_2, E_3; therefore, these four elements generate a U(2) subgroup of SU(3). This U(2) subgroup, often sloppily labeled as SU(2)xU(1), is given by those 3 by 3 matrices that contain a unitary 2 by 2 matrix in the upper left, contain zeroes in the remaining four off-diagonal elements, and contain the inverse value of the determinant of the 2 by 2 matrix in the remaining, lower right diagonal element. In short, SU(3) contains three linearly independent U(2) subgroups.

SU(3) is characterized by the way that the SU(2) triplets are connected. In particular, the product $E_3E_9E_{10}$ is the identity, reflecting the linear dependence of the three corresponding generators λ_n. We also have $E_8E_9 = E_{10}$. Also the product of E_8 with its

TABLE 10 The multiplication table for the generators λ_1 to λ_8 of SU(3), and for the additional, *linearly dependent* matrices $\lambda_9 = -\lambda_3/2 - \lambda_8\sqrt{3}/2$ and $\lambda_{10} = -\lambda_3/2 + \lambda_8\sqrt{3}/2$ that are *not* generators. Note that, despite the appearance, $\lambda_4^2 = \lambda_5^2 = \lambda_9^2$ and $\lambda_6^2 = \lambda_7^2 = \lambda_{10}^2$.

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_9	λ_6	λ_7	λ_{10}	λ_8
λ_1	$2/3$ $+\lambda_8/\sqrt{3}$	$i\lambda_3$	$-i\lambda_2$	$\lambda_6/2$ $+i\lambda_7/2$	$-i\lambda_6/2$ $+\lambda_7/2$	$-\lambda_1/2$ $+i\lambda_2/2$	$\lambda_4/2$ $+i\lambda_5/2$	$-i\lambda_4/2$ $+\lambda_5/2$	$\lambda_1/2$ $+i\lambda_2/2$	$\lambda_1/\sqrt{3}$
λ_2	$-i\lambda_3$	$2/3$ $+\lambda_8/\sqrt{3}$	$i\lambda_1$	$i\lambda_6/2$ $-\lambda_7/2$	$\lambda_6/2$ $+i\lambda_7/2$	$-i\lambda_1/2$ $-\lambda_2/2$	$-i\lambda_4/2$ $+\lambda_5/2$	$-\lambda_4/2$ $-i\lambda_5/2$	$-i\lambda_1/2$ $+\lambda_2/2$	$\lambda_2/\sqrt{3}$
λ_3	$i\lambda_2$	$-i\lambda_1$	$2/3$ $+\lambda_8/\sqrt{3}$	$\lambda_4/2$ $+i\lambda_5/2$	$-i\lambda_4/2$ $+\lambda_5/2$	$-1/3-\lambda_3/3$ $+\lambda_9/3$	$-\lambda_6/2$ $-i\lambda_7/2$	$i\lambda_6/2$ $-\lambda_7/2$	$-1/3+\lambda_3/3$ $+\lambda_{10}/3$	$\lambda_3/\sqrt{3}$
λ_4	$\lambda_6/2$ $-i\lambda_7/2$	$-i\lambda_6/2$ $-\lambda_7/2$	$\lambda_4/2$ $-i\lambda_5/2$	$2/3+\lambda_3/2$ $-\lambda_8/2\sqrt{3}$	$-i\lambda_9$	$i\lambda_5$	$\lambda_1/2$ $+i\lambda_2/2$	$i\lambda_1/2$ $-\lambda_2/2$	$-\lambda_4/2$ $-i\lambda_5/2$	$-\lambda_4/2\sqrt{3}$ $-i\sqrt{3}\lambda_5/2$
λ_5	$i\lambda_6/2$ $+\lambda_7/2$	$\lambda_6/2$ $-i\lambda_7/2$	$i\lambda_4/2$ $+\lambda_5/2$	$i\lambda_9$	$2/3+\lambda_3/2$ $-\lambda_8/2\sqrt{3}$	$-i\lambda_4$	$-i\lambda_1/2$ $+\lambda_2/2$	$\lambda_1/2$ $+i\lambda_2/2$	$i\lambda_4/2$ $-\lambda_5/2$	$i\sqrt{3}\lambda_4/2$ $-\lambda_5/2\sqrt{3}$
λ_9	$-\lambda_1/2$ $-i\lambda_2/2$	$i\lambda_1/2$ $-\lambda_2/2$	$-1/3-\lambda_3/3$ $+\lambda_9/3$	$-i\lambda_5$	$i\lambda_4$	$2/3+2\lambda_3/3$ $+\lambda_9/3$	$\lambda_6/2$ $-i\lambda_7/2$	$i\lambda_6/2$ $+\lambda_7/2$	$-1/3-\lambda_9/3$ $+\lambda_{10}/3$	-1 $+\lambda_{10}$
λ_6	$+\lambda_4/2$ $-i\lambda_5/2$	$i\lambda_4/2$ $+\lambda_5/2$	$-\lambda_6/2$ $+i\lambda_7/2$	$\lambda_1/2$ $-i\lambda_2/2$	$i\lambda_1/2$ $+\lambda_2/2$	$\lambda_6/2$ $+i\lambda_7/2$	$2/3-\lambda_3/2$ $-\lambda_8/2\sqrt{3}$	$i\lambda_{10}$	$-i\lambda_7$	$-\lambda_6/2\sqrt{3}$ $-i\sqrt{3}\lambda_7/2$
λ_7	$i\lambda_4/2$ $+\lambda_5/2$	$-\lambda_4/2$ $+i\lambda_5/2$	$-i\lambda_6/2$ $-\lambda_7/2$	$-i\lambda_1/2$ $-\lambda_2/2$	$\lambda_1/2$ $-i\lambda_2/2$	$-i\lambda_6/2$ $+\lambda_7/2$	$-i\lambda_{10}$	$2/3-\lambda_3/2$ $-\lambda_8/2\sqrt{3}$	$i\lambda_6$	$i\sqrt{3}\lambda_6/2$ $-\lambda_7/2\sqrt{3}$
λ_{10}	$-\lambda_1/2$ $+i\lambda_2/2$	$-i\lambda_1/2$ $-\lambda_2/2$	$-1/3+\lambda_3/3$ $-\lambda_{10}/3$	$-\lambda_4/2$ $+i\lambda_5/2$	$-i\lambda_4/2$ $-\lambda_5/2$	$-1/3-\lambda_9/3$ $+\lambda_{10}/3$	$i\lambda_7$	$-i\lambda_6$	$2/3-\lambda_3/3$ $+\lambda_9/3$	1 $+\lambda_9$
λ_8	$\lambda_1/\sqrt{3}$	$\lambda_2/\sqrt{3}$	$\lambda_3/\sqrt{3}$	$-\lambda_4/2\sqrt{3}$ $+i\sqrt{3}\lambda_5/2$	$-i\sqrt{3}\lambda_4/2$ $-\lambda_5/2\sqrt{3}$	-1 $+\lambda_{10}$	$-\lambda_6/2\sqrt{3}$ $+i\sqrt{3}\lambda_7/2$	$-i\sqrt{3}\lambda_6/2$ $-\lambda_7/2\sqrt{3}$	1 $+\lambda_9$	$2/3$ $-\lambda_8/\sqrt{3}$

companions from the other two triplets is the identity.

Finally, the product $(E_k E_l)^3$ for any k taken from the set $(1, 2, 4, 5, 6, 7)$ and any l from the same set, but from a *different* triplet, is also the identity matrix. This property of the third powers – taken together with the threefold symmetry of its centre – can be seen as the essential property that distinguishes SU(3) from other Lie groups. We now return to the strand model and show that slides indeed define an SU(3) group.

From slides to SU(3)

The *slide*, or *third Reidemeister move*, involves *three* pieces of strands. The textbook version of the third Reidemeister move – which is called E_0 here and is illustrated in Figure 69 – moves or 'slides' one strand, taken to be the horizontal blue one in the figure,

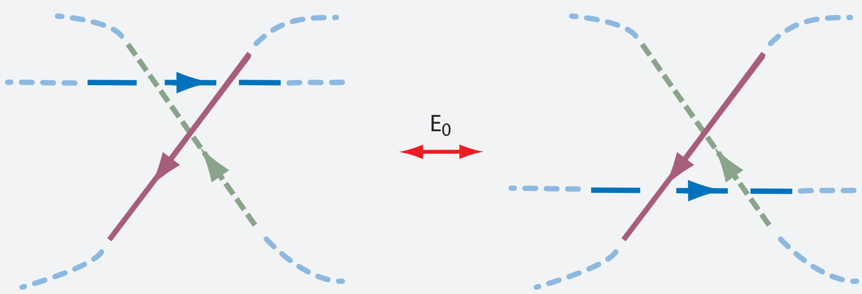


FIGURE 69 The textbook version E_0 of the slide move, or third Reidemeister move, is unobservable, because it does not involve crossing switches.

against a crossing of the other two. Equivalently, we can say that a slide pushes two strands against the blue strand that is kept in place. This textbook slide – we also call it a *pure slide* here – does not contain any crossing switch; following the fundamental principle of the strand model, it is therefore unobservable, or, simply said, of no physical relevance. However, related strand moves that do involve crossing switches do exist.

We introduce eight *generalized slides*, or slide-rotations, for a three-strand configuration; they are shown in Figure 70. We directly call these generalized slides E_1 to E_8, because they will turn out to correspond to the SU(3) group elements with the same name that were introduced above. In other words, we will show that the *generalized slides* E_n are elements of a Lie group SU(3); in particular, they obey all the properties expected from the correspondence with the SU(3) generators λ_n in Gell-Mann's choice:

$$E_n = \mathrm{e}^{\pi \mathrm{i} \lambda_n / 2} . \tag{164}$$

In the strand model, the generators λ_n describe the difference between an infinitesimal generalized slide – thus a slide-rotation with a rotation by an infinitesimal angle – and the identity. For slides, concatenation is equivalent to group multiplication, as expected. Slides form a group. We will now show that the slide generators obey the multiplication table already given in Table 10.

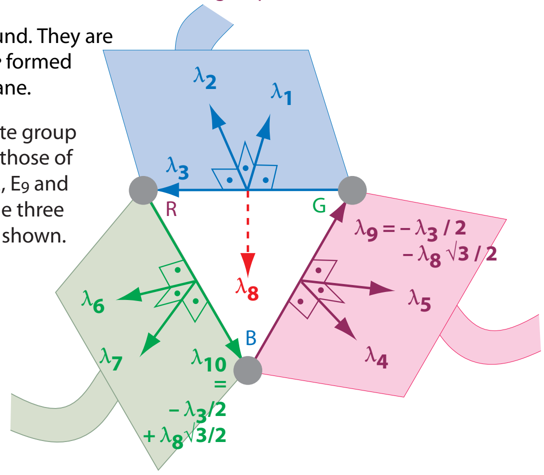
To see how the SU(3) multiplication table follows from Figure 70, we first note that the starting strand configuration of the Reidemeister III move contains, if all spatial configurations are considered, the same threefold symmetry as the centre of SU(3). In particular, like the generators and the basis vectors of SU(3), also the slides of the figure can be grouped into three triplets.

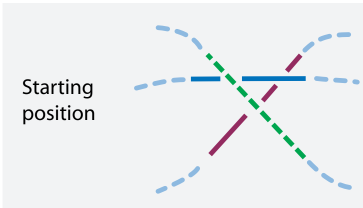
We now focus on the first triplet, the one formed by the three slides E_1, E_2 and E_3.

The *generalized slides*, or *Reidemeister III moves*, acting on three strands, form an SU(3) group.

The 8 generalized slides are shown below, with grey background. They are local slides and rotations by an angle π of an *imaginary buckle* formed by (usually) two strands. The strands lie (mostly) in a paper plane.

For each SU(2) triplet inside SU(3), the rotation axes of the finite group elements E_n are arranged at right angles to each other, as are those of the generators λ_n shown on the right. The rotation axes for E_3, E_9 and E_{10} are parallel; they are perpendicular to the paper plane. The three imaginary belt buckles for the three SU(2) subgroups are also shown.





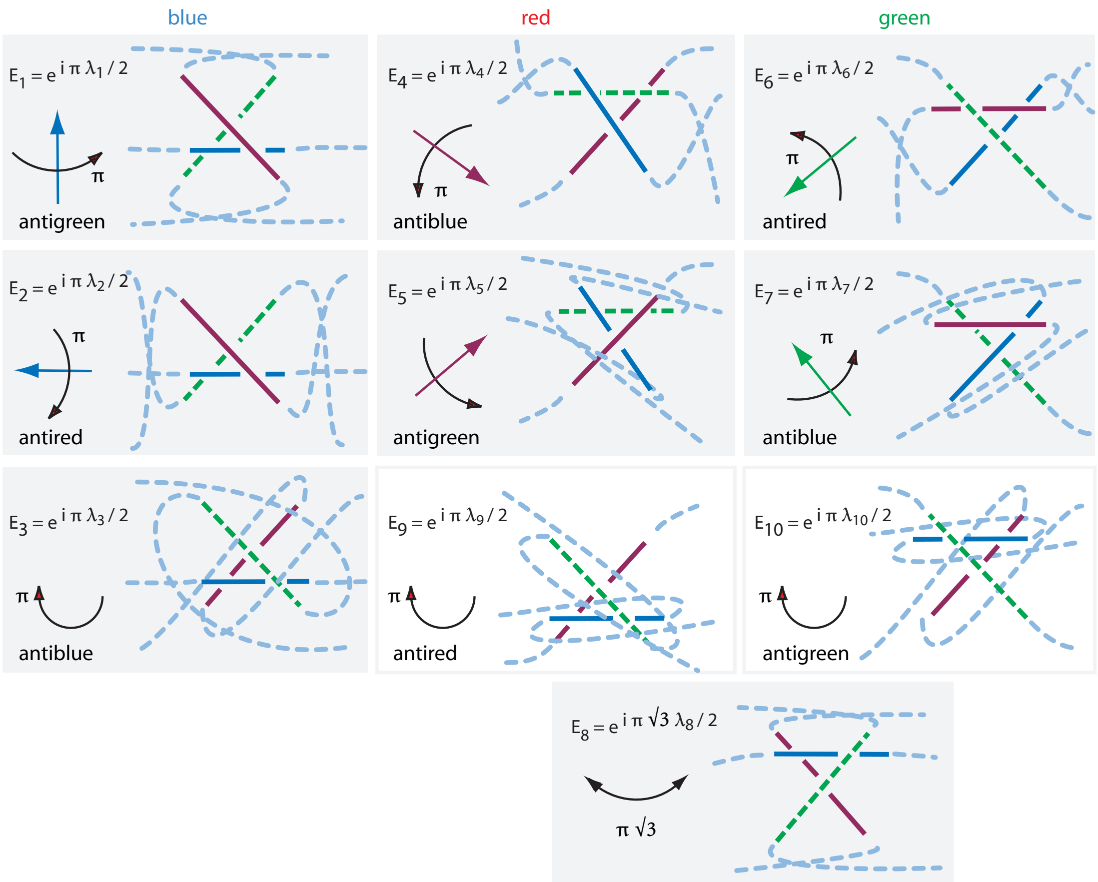


FIGURE 70 The strand deformations for the generalized slide moves E_n. The corresponding generators λ_n lead to an SU(3) structure, as shown in the text. Note that the rotation vectors for the generators λ_n and for the generalized slide moves E_n differ from each other. For clarity, the figure shows, instead of the deformation of the strand under discussion, the complementary deformations of the other two strands.

To makes things clear, these moves can be pictured as combined deformations and slides of the red and green strands against the horizontal blue strand. We can imagine these moves like those of the belt trick, but acting on an *imaginary buckle* formed only by the red and green strands. These generalized slides do contain crossing changes; therefore they are observable and are of physical relevance.

We note that 'slide' is not a perfect term for the generalized deformations E_1 to E_8; in fact, we might prefer to call them *slide-rotations*, because they are slide-rotations by an angle π that are applied to an imaginary belt buckle. Despite the involved construction, these generalized, observable moves remain modelled on the textbook slide E_0; in particular, they require *three* strand segments. The generalized, observable moves just defined *differ* from the twists and pokes discussed above, in the sections on the electromagnetic and weak interactions; thus they differ from Reidemeister I and II moves. As a result, we will usually continue to call the generalized, observable moves simply *slides*.

For simplicity, we assume – similarly to what we did in the discussion about the weak interaction – that the three strand segments are (roughly) in a plane. This is an idealized situation; in fact, the arguments given in the following apply also to all other three-dimensional configurations of three strands. In particular, the same results appear if all three strands segments are assumed perpendicular to each other, instead of lying in a plane.

We note that the rotation axes of the generalized slides E_1 and E_2 are neither aligned nor orthogonal to the paper plane. More precisely, the rotation axes of E_1, E_4 and E_6 are perpendicular to the sides of a cube. E_2, E_5 and E_7 are perpendicular to them. For the first triplet, the rotation axes E_1, E_2 and E_3 form an orthonormal basis; the same is valid for the other two triplets. We now show that the slides of the first triplet define an SU(2) group.

The observable, generalized slides in the triplet E_1, E_2 and E_3 can be concatenated. We distinguish two cases. The first case is the concatenation of any such slide with itself. The result corresponds to a rotation by 2π of the chosen strand pair and its imaginary belt buckle, and thus induces a corresponding amount of tail twisting. In fact, when any slide of the triplet is concatenated *four* times with itself, the result is the identity operation. Comparing a twofold and a fourfold concatenation, we see that they differ only by an entangling, or algebraically, by a minus sign for the imaginary buckle. This already realizes half of the belt trick that visualizes SU(2).

The other case to be checked is the concatenation of two different slides of the triplet. The result is always the third slide of the triplet (up to a sign that depends on whether the combination is cyclical or not). This behaviour realizes the other half of the belt trick. In short, we have shown that the triplet containing the first three generalized slides defines an SU(2) group. More precisely, the infinitesimal slide-rotations λ_1, λ_2 and λ_3 corresponding to the finite SU(3) elements E_1, E_2 and E_3 generate the SU(2) Lie algebra of an SU(2) Lie group. The SU(2) subgroup just found is just one of the three linearly independent SU(2) subgroups of SU(3). The generators of the first slide triplet thus reproduce the nine results in the upper left of Table 10. We thus retain that we can indeed visualize the first three generalized slides with the help of the three orthogonal rotations by π of an imaginary belt buckle formed by the red and green strands.

Page 263

For the visualization of SU(3) it is essential to recall that the direction in three-dimensional space of the vectors visualizing λ_n and those visualizing E_n differ from each other. This already the case for U(1).

The remaining generalized slides that are possible in the three-strand configuration are easily constructed using the threefold symmetry of the strand configuration; they are illustrated in Figure 70. For each of the three strand segments there is a triplet of observable slides; this yields a total of *nine* possible generalized slides for the observer defined by the paper plane. In the second triplet, the slides corresponding to E_1 and E_2 are called E_4 and E_5, and in the third triplet they are called E_6 and E_7. For the three slides corresponding to E_3 – we call the other two E_9 and E_{10} – *only two* generators are linearly independent. Indeed, the figure shows that $E_3E_9E_{10}$ – whose axes are all three parallel – is the identity matrix; this expected from an SU(3) structure. The three operations E_3, E_9 and E_{10} also commute with all other operations; thus they form the centre of the group defined by all E. The second linearly independent, generalized slide of common use, E_8, is also shown in the figure; it is a linear combination of E_9 and E_{10}. We note that the strand model also visualizes the factor $\sqrt{3}$ in the definition of E_8. In total, we get *eight* linearly independent generalized slides. All slides, except for E_8, act on an imaginary belt buckle that is formed by two strands.

We saw that the generators corresponding to the slides E_1, E_2 and E_3 generate an SU(2) subgroup. The same holds for the corresponding triplet E_4, E_5 and the linear combination $E_9 = -E_3/2 - E_8\sqrt{3}/2$ (corresponding to E_3), and for the triplet E_6, E_7 and $E_{10} = -E_3/2 + E_8\sqrt{3}/2$. For each of these slides, a fourfold concatenation yields the identity; and inside each triplet, the concatenation of two different slides yields a multiple of the third slide. In short, for each triplet, the corresponding infinitesimal slides generate an SU(2) group. These three SU(2) groups are linearly independent. We have thus reproduced an important part of the structure of SU(3). In addition, we have found a visualization of SU(3); since each SU(2) group can be represented by a separate imaginary buckle, the group SU(3) can be visualized – in many, but not all in aspects – with the help of three imaginary buckles. The top right of Figure 70 illustrates this visualization.

The correspondence of the slides and the multiplication table increases further if we change slightly the definition of the first triplet. In this first triplet we can take as imaginary buckle the set of *all three* central segments. Moving all three strands together simplifies the visualization, because for the first triplert, the blue strand is trapped between the other two strands. In this way, generalized slide still consists of a rotation followed by a slide. And we still have a SU(2) subgroup for the first triplet.

The slide E_8 differs from the other slides, as expected from SU(3). It describes a motion that rotates the red and green strands in opposite directions; this is illustrated in Figure 70. E_8 is thus *not* well described with an imaginary belt buckle. It is straightforward to check that the slide E_8 commutes with E_1, E_2, E_3 and obviously with itself, but not with the other generalized slides. Together, E_8 and the first triplet thus form a U(2) Lie group, as expected. In addition, we find that E_8 commutes with E_9 and E_{10}, and that $E_8E_9 = E_{10}$, as expected from SU(3). The strand model also implies that the product of E_8 with its two counterparts from the other triplets is the identity matrix, as expected from SU(3).

The last step to show the equivalence of slides and SU(3) requires us to confirm the multiplication properties – between slides E_n or between generators λ_n – from *different* triplets. In fact, because of the three-fold symmetry of the centre, we only need to check two multiplication results between slides from different triplets: one that either involves

λ_3 or λ_8, and one that does not.

We begin with products involving λ_3 and one of the first two elements of another triplet. Such products yield a weighted sum of generators of the triplet. It is easier to check these product properties by using the exemplary relation between finite group elements $E_5 E_3 E_4 = E_3$. Note that only this specific permutation of 5, 3 and 4 yields this result. Playing with the strand model confirms the relation. Similar comments apply to $E_6 E_3 E_7 = E_3$ – and to the corresponding products involving E_9, such as $E_1 E_9 E_2 = E_9$, or E_{10}, such as $E_1 E_{10} E_2 = E_{10}$ – as well as $E_4 E_8 E_5 = E_8$ and $E_6 E_8 E_7 = E_8$. The strand model allows anybody to check that these relations are satisfied.

We continue with the exemplary product $\lambda_5 \lambda_7$, respectively $E_5 E_7$. We note a basic difference between a product like $\lambda_5 \lambda_7$ and any product of two generators from the same triplet. The product $\lambda_5 \lambda_7$ – like the other concatenations of generators from different triplets – does not yield a single generator, but yields a combination, i.e., a *sum* of generators. The combination is not easy to visualize with strands; an easier way is to check the SU(3) algebra using the properties of the product $E_5 E_7$.

As mentioned above, in SU(3), for products involving the first two members from different triplets, *the threefold concatenation* $(E_i E_j)^3$ *is the identity.* And indeed, Figure 70 confirms that $(E_2 E_4)^3$ or $(E_5 E_7)^3$ is the identity. Similarly, also the other products can be tested with the help of three strands.

Using the visualization with three strands, we have thus confirmed all products of generators from two different triplets that appear in Table 10. We note that Figure 70 also illustrates that the three slides E_2, E_5 and E_7 generate an SO(3) group, the rotation group in three dimensions. In order to see this, we observe that the infinitesimal versions of the three slides generate all possible rotations in three dimensions of the central triangle. An SO(3) group also appears for the slides 1, 4 and 7, for the slides 1, 5 and 6, and for the slides 2, 4 and 6. These are the four basic SO(3) subgroups of SU(3). The remaining combinations of three operations from three different triplets – such as 1, 4 and 6, or the combination 1, 5 and 7, or the combination 2, 4 and 7, or the combination 2, 5 and 6 – do *not* generate any subgroup. This can be confirmed by exploring the corresponding strand moves.

We can conclude: in a region with three strands crossing each other, the eight linearly independent, generalized slides that can be applied to that region define the group SU(3). In other words, *the group SU(3) follows from the third Reidemeister move.*

In the same way as for the other gauge groups, we find that particles whose strand models contain configurations with three strand segments can be subject to an SU(3) gauge interaction. In experiments, this interaction is called the *strong nuclear interaction*. The strong interaction is due to the Reidemeister III move. Like for the other interactions, a particle will only interact strongly if its tangle is not too symmetric, because in the symmetric case, averaged over time, there will be no net interaction. We will clarify the details below, when we discuss the specific tangles and colour charges of the different elementary matter particles.

Page 313

The strand model for gluons

Physically, the eight slides corresponding to the Gell-Mann matrices represent the effects of the eight *gluons*, the intermediate vector bosons of the strong interaction, that can act on a particle.

▷ Given that the eight slides E_1 to E_8 represent the effects of the eight gluons, they also represent the gluons themselves.

Interactions are transfers of a tangle process to another tangle. Therefore

▷ The absorption of a gluon is a slide that is transferred to another particle.
▷ The emission of a gluon is a slide that is transferred to three vacuum strands.

To visualize the concept of gluon even further, we can say that every gluon can be described as a strand structure that continuously performs an SU(3) operation, i.e., a generalized slide continuously repeating itself. We found a similar correspondence for the other gauge interactions. In case of the electromagnetic interaction, the intermediate vector boson, the photon, can be described as a strand that continuously performs a U(1) operation, i.e., a rotation. In case of the weak interaction, a weak intermediate vector boson can be described as a strand that continuously performs an SU(2) operation, i.e., an operation from the belt trick. This is most evident in the unbroken form of the weak bosons.

Challenge 168 e

Every gluon can also be seen as the deformation of a single strand that drags its surrounding with it. This single strand description of gluons implies that gluons have vanishing mass and vanishing charge. This single strand description of gluons also implies that they have spin 1, as is observed. The strand model of the gluon also implies that free gluons would have a huge energy.

The SU(3) multiplication table confirms that the eight gluons transform according to the adjoint (and faithful) representation of SU(3). Therefore, each row or column in a Gell-Mann matrix thus corresponds to one of the three *colours* of the strong interaction. The exploration of slide concatenation also showed that two general slides do not commute and do not anticommute. The group SU(3) is *non-Abelian*. This implies that gluons interact among themselves. Both the multiplication table and the strand model for gluons imply that two interacting gluons can yield either one or two new gluons, but not more. This is illustrated in Figure 71. The strand model, through its generation of SU(3), thus implies that gluons interact among themselves, but only in triple and quartic gluon vertices.

Slides – i.e., gluon emission or absorption – never change the topology of tangles, and in particular, of matter tangles. Therefore, the strand model predicts that the strong interactions conserve electric charge, baryon number, weak isospin, flavour, spin and all parities. This is indeed observed. In particular, there is a natural lack of C, P and CP violation by slides. This is precisely what is observed for the strong interaction.

Because gluons do not change the topology of the particle tangles they act upon, but only change their shape, gluons are predicted to be massless in the strand model, despite interacting among themselves. And because gluons interact among themselves, free gluons are predicted not to appear in nature. And of course, all these conclusions agree with experiments.

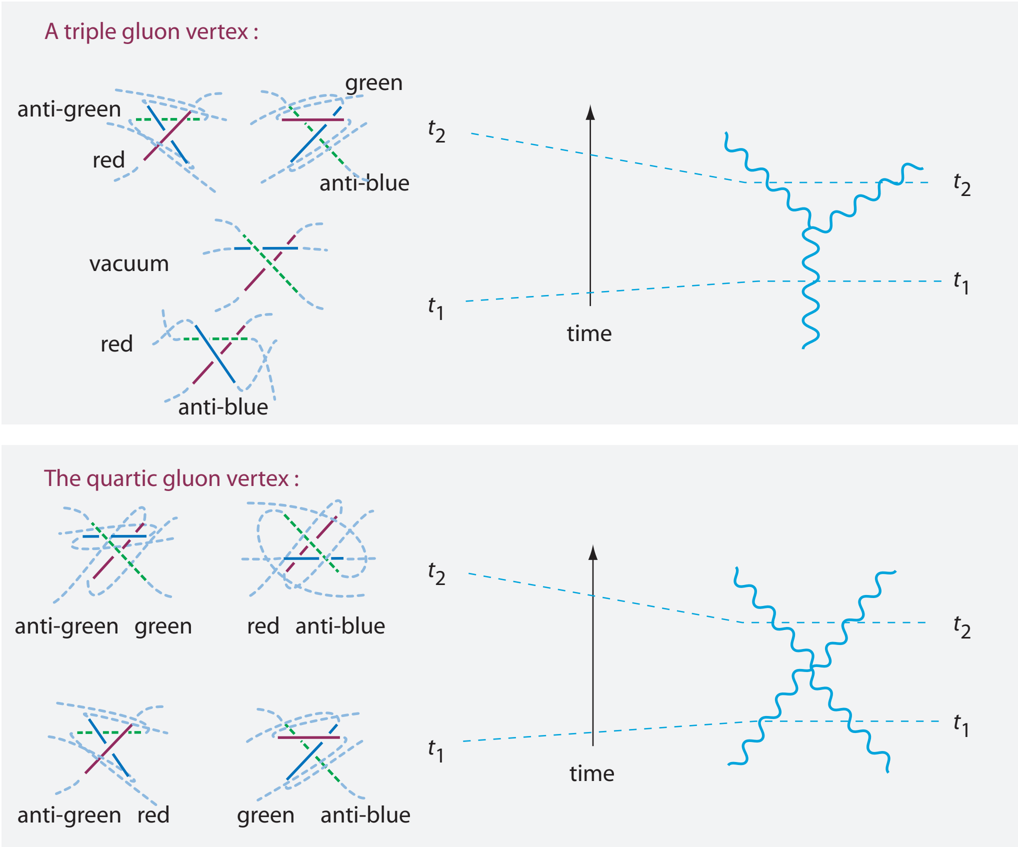


FIGURE 71 The two types of self-interaction of gluons in the strand model.

In summary, we have shown that in the strand model, the strong nuclear interaction and all its properties appear automatically form slides, i.e., from Reidemeister III moves. In particular, the strand model implies that the Lagrangian of strongly interacting fermions has a SU(3) gauge invariance that is due to generalized slide deformations.

The gluon Lagrangian

Gluons are massless particles with spin 1. As a result, the field intensities and the Lagrangian are determined in the same way as for photons: energy density is the square of crossing density, i.e., the 'square' of field intensity. Since there are 8 gluons, the Lagrangian density becomes

$$\mathcal{L}_{\text{gluons}} = -\frac{1}{4}\sum_{a=1}^{8} G^a_{\mu\nu} G_a^{\mu\nu} \tag{165}$$

where the gluon field intensities, with *two* greek indices, are given naturally as

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g f^{abc} G^b_\mu G^c_\nu \ , \tag{166}$$

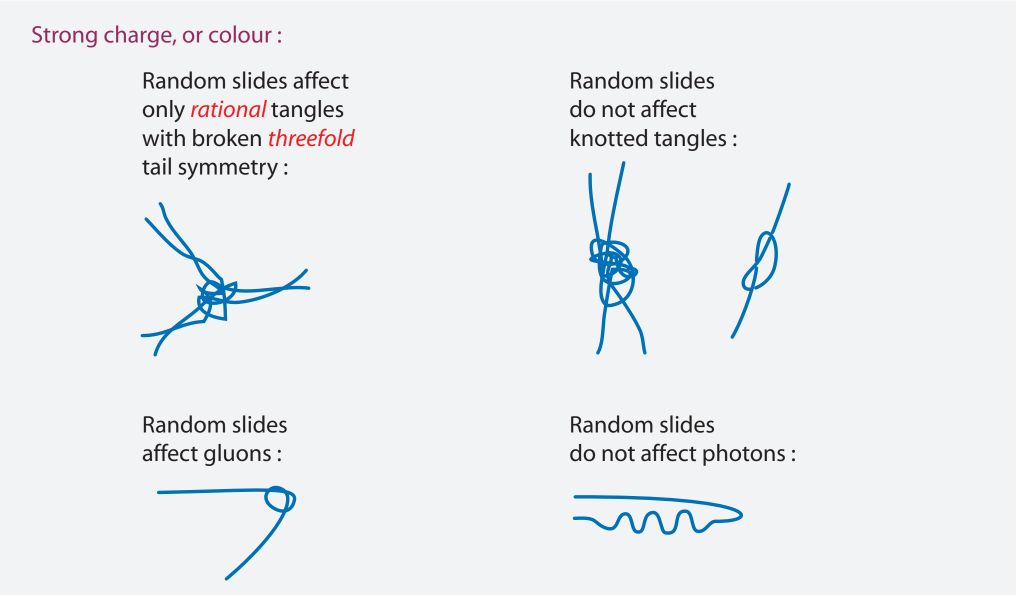


FIGURE 72 Tangles with and without colour charge. (This image needs to be updated: no knotted tangles occur in nature.)

and f_{abc} are the structure constants of SU(3) that can be deduced from the multiplication table given above. The quantities G^a_μ, with *one* greek index, are the gluon vector potentials. The last term in the definition of the field intensities corresponds to the triple and quartic vertices in the Feynman diagrams of gluon interactions. They are shown in Figure 71. The Lagrangian is simply the natural generalization from the U(1) case of photons to the SU(3) case of gluons. In short, we obtain the usual free gluon Lagrangian from the strand model. Page 234

Colour charge

Surrounded by a bath of gluons that randomly induce slides of all kinds, not all fermion cores will change their rotation state. Generally speaking, particles have colour if a bath of random gluons changes their phase. Only tangles which lack some symmetry will therefore possess colour charge. Tangle that are symmetric will be neutral, or 'white'. Which symmetry is important here?

We see directly that the *photon* tangle is not sensitive to a gluon bath. The same is valid for W and Z bosons. These tangles are too simple. The strand model predicts that these particles are colour-neutral, i.e., that they are 'white', as is observed.

On the other hand, the multiplication properties given above shows that *gluons* interact among themselves and thus that they have colour charge. In fact, group theory shows that their properties are best described by saying that gluons have a colour and an anticolour; this is the simplest way to describe the representation to which they belong. In short, the strand model of gluons automatically implies that they carry both a colour and an anti-colour.

Fermions behave differently. In the strand model, a fermion has colour charge if the

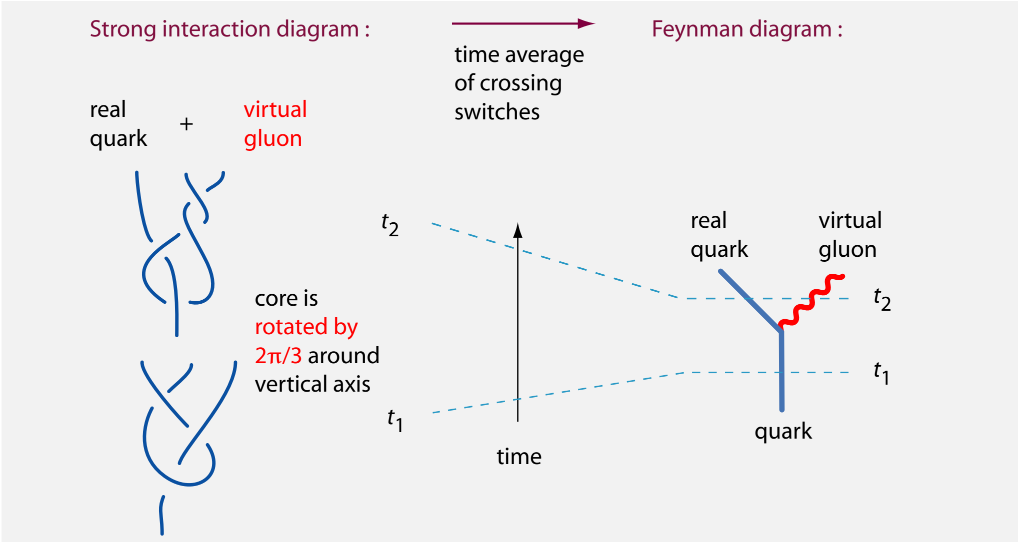


FIGURE 73 The Feynman diagram of the strong interaction for a quark. The upper triplet of tails correspond to the three belts.

corresponding triple belt model is affected by large numbers of random gluons. The first tangles that come to mind are tangles made of three strands, such as the simple tangles shown in Figure 70. But a short investigation shows that such tangles are colour-neutral, or 'white'. We will see below that this implies that *leptons* are colour-neutral, or 'white'. In contrast, a rational fermion tangle does not suffer this fate. (We recall that a so-called *rational tangle* is by definition made of exactly *two* strands; a two-stranded tangle is rational if the two strands can be untangled just by moving the tails around.) In a bath of gluon strands that induce slides, i.e., third Reidemeister moves, a general *rational tangle* made of two strands is expected to be influenced, and thus to be colour-charged.

Rational tangles made of two strands are the simplest possible tangles with colour. A tangle is called *rational* if it can be untangled just by moving the tails around. An example of a rational tangle is shown in Figure 73. Such tangles break the three-fold symmetry of the three-belt structure, and are thus colour-charged. We will show below how these tangles are related to *quarks*. We can thus say:

▷ A fermion tangle has ***colour charge*** if its three-belt model is not symmetric for rotations by $\pm 2\pi/3$.

Coloured rational tangles automatically have *three* possible colours:

▷ The ***three colour charges*** are the three possibilities to map a tangle to the three belt model.* *Each colour is thus a particular orientation in ordinary space.*

If we want to explore more complicated types of tangles of two strands, such as *prime*

Challenge 169 ny

* Can you define a geometric or even a topological knot invariant that reproduces colour charge?

Page 318 tangles or *locally knotted* tangles, we recall that such tangles are not part of the strand model. The strand model thus predicts that rational tangles made of two strands are the basic colour states. And indeed, in nature, quarks are the only fermions with colour charge.

We can summarize that colour charge is related to orientation in space. The three possible colours and anticolours are consequences of the possible orientations along the three dimensions of space.

Properties of the strong interaction

In the strand model, all interactions are *deformations* of the tangle core. Specifically, the strong interaction is due to exchange of *slides*. Particles have strong charge, or colour, if their tangles lack the three-belt symmetry just specified. In the case of coloured fermions, *colour change* is a change of the mapping to the three-belt model, i.e., a change of orientation of the tangle in space.

If we use the strand definition of the strong interaction, visual inspection shows us that slide exchanges, and thus gluon exchanges, are deformations that conserve topology; therefore gluon exchange *conserves colour*. Since the strong interaction conserves the topology of all involved tangles and knots, the strong interaction also *conserves electric charge*, *parity*, and, as we shall see below, *all other quantum numbers* – except colour itself, of course. All these results correspond to observation.

The Lagrangian of QCD

We started from the idea that tangle core deformations lead to phase redefinitions. We then found that slides imply that the strong interaction Lagrangian for matter and for radiation fields is SU(3) gauge invariant. If we include these two gauge invariances into the fermion Lagrangian density from the Dirac equation, we get

$$\mathcal{L}_{\text{QCD}} = \sum_q \overline{\Psi}_q (i\hbar c \not{D} - m_q c^2 \delta_{qq'}) \Psi_{q'} - \frac{1}{4} \sum_{a=1}^{8} G^a{}_{\mu\nu} G_a{}^{\mu\nu} , \qquad (167)$$

where the index q counts the coloured fermion, i.e., the quark. In this Lagrangian density, $\not{D}$ is now the SU(3) gauge covariant derivative

$$\not{D} = \not{\partial} - g\, \gamma^\mu G^a_\mu \lambda_a , \qquad (168)$$

where g is the gauge coupling, λ_a are the generators of SU(3), i.e., the Gell-Mann matrices given above, and the G^a_μ are, as before, the gluon vector potentials. The last term in the covariant derivative corresponds to the Feynman diagram and the strand diagram of Figure 73. This is the Lagrangian density of QCD.

In summary: the strand model reproduces QCD. However, we have not yet deduced the number and masses m_q of the quarks, nor the strong gauge coupling g.

Renormalization of the strong interaction

The slide move description of the strong interaction implies that only three Feynman diagrams are possible: one QCD Feynman diagram is possible for quarks, and only the triple and the quartic vertices are possible among gluons. This limited range of options allowed us to deduce the QCD Lagrangian. The limited range of options is also essential for the *renormalization* of QCD. The strand model thus automatically ensures that the strong interaction is renormalizable.

In short, the strand model provides a new underlying picture for the Feynman diagrams of the strong interaction, but does not change the physical results at any energy scale accessible in the laboratory. In particular, the measured running of the strong coupling constant is reproduced. Indeed, in the strand model, a flux-tube–like bond between the quarks appears automatically, as we will see when exploring hadrons. At high kinetic energies, the bond has little effect, so that quarks behave more like free particles. In short, we find that the strand model reproduces *asymptotic freedom* and also provides an argument for quark confinement. We will return to the issue in more detail below.

Page 386
Page 332, page 340
Page 337

Curiosities and fun challenges about SU(3)

Vol. V, page 289

Deducing the Lie group SU(3) from a three-dimensional model is a new result. In particular, deducing the gauge group SU(3) as a *deformation gauge group* is new. Frank Wilczek, Alfred Shapere, Alden Mead, Jerry Marsden and several others have confirmed that before this discovery, only the geometric Lie group SO(3) and its subgroups had been found in deformations. The fundamental principle of the strand model shows its power by overcoming this limitation. (Apparently, nobody had even realized that the belt trick already implies the possibility of an SU(2) gauge group for deformations.)

Ref. 186

* *

We have discussed the *shape deformations* that lead to the SU(3) group. But what are the precise *phase choices* for a crossing that lead to SU(3) invariance?

Challenge 170 ny

* *

Do the two linear independent gluons with lined-up tails have the same properties as the other six gluons?

Challenge 171 ny

* *

Three strands can cross each other also in another way, such that the three strands are interlocked. Why can we disregard the situation in this section?

Challenge 172 s

* *

Deducing the Lie groups U(1), SU(2) and SU(3) directly from a basic principle contradicts another old dream. Many scholars hoped that the three gauge groups have something to do with the sequence complex numbers, quaternions and octonions. The strand model quashes this hope – or at least changes it in an almost unrecognizable way.

* *

Challenge 173 e

The tangles for the W and Z bosons have no colour charge. Can you confirm this?

* *

Challenge 174 ny The Lie group SU(3) is also the symmetry group of the three-dimensional harmonic oscillator. What is the geometric relation to the Lie group SU(3) induced by slides?

* *

Challenge 175 e Confirm that the strand model does not contradict the Coleman–Mandula theorem on the possible conserved quantities in quantum field theory.

* *

Challenge 176 e Confirm that the strand model does not contradict the Weinberg–Witten theorem on the possible massless particles in quantum field theory.

* *

Challenge 177 d Are the *Wightman axioms* of quantum field theory fulfilled by the strand model with interactions? The *Haag–Kastler axioms*? Is Haag's theorem circumvented?

* *

Ref. 187 Challenge 178 ny Show that the BCFW recursion relation for tree level gluon scattering follows from the strand model.

Summary on the strong interaction and experimental predictions

We have deduced the Lagrangian density of QCD from the strand model with the help of slides. Is there a difference between the strand model and QCD? No, not as long as gravity plays no role. The strand model predicts that gravitation only comes into play near the Planck energy $\sqrt{\hbar c^5/4G}$. And indeed, accelerator experiments have not yet found any effect that contradicts QCD, and therefore no effect that contradicts the strand model of the strong interaction.

The strand model also predicts that the strong interaction is naturally CP-invariant. This means that axions – particles invented to explain the invariance – are unnecessary: Page 353 as shown below, the strand model even predicts that they do not to exist. Both predictions agree with experiment.

The strand model of the strong interaction implies that the SU(3) gauge symmetry is valid at all energies. No other gauge group plays a role in the strong interaction. The strand model thus predicts again that there is no grand unification in nature, and thus no larger gauge group. Often discussed groups such as SU(5), SO(10), E6, E7, E8 or SO(32) are predicted not to apply to nature. Also this prediction is not contradicted by experiment.

The strand model further predicts that the combination of gravity and quantum theory turns all Planck units into *limit* values. The strand model thus predicts a maximum strong field value given by the Planck force divided by the strong charge of the quark. All physical systems – including all astrophysical objects, such as neutron stars, quark stars, gamma-ray bursters or quasars – are predicted to conform to this field limit. So far, this prediction is validated by experiment.

In summary, we have shown that Reidemeister III moves – or slides – in tangle cores lead to an SU(3) gauge invariance and a Lagrangian that reproduces the strong interaction. Colour charge is related to the topology of certain rational tangles. In this way, we have deduced the origin and most observed properties of the strong interaction. We have thus settled another issue of the millennium list. However, we still need to deduce the tangles and the number of quarks, their masses and the strength of the strong coupling.

Page 164

Summary and predictions about gauge interactions

At this point of our adventure, we have deduced gauge theory and the three known gauge interactions from strands. Using only the fundamental principle, we explained the dimensions of space-time, the Planck units, the principle of least action, the appearance of the gauge groups U(1), broken SU(2) and SU(3), of renormalization, of Lorentz symmetry and of permutation symmetry. Thus we have deduced all the concepts and all the mathematical structures that are necessary to *formulate* the standard model of elementary particles.

Page 152

In particular, the strand model provides a description and explanation of the three gauge interactions at Planck scales that is based on *deformations* of strands. The description of QED is illustrated in Figure 74: it shows the way that strands model the emission of a photon by an electron. The deduction of the three gauge interactions given in this text, with the help of the Reidemeister moves, is the first and, at present, the *only* explanation of the three gauge forces. No other explanation or deduction has ever been given.

We have shown that quantum field theory is an *approximation* of the strand model. The approximation appears when the strand diameter is neglected; quantum field theory is thus valid for all energies below the Planck scale. In other words, in contrast to many other attempts at unification, the strand model is *not a generalization* of quantum field theory. The strand model for the three gauge interactions is also unmodifiable. These properties are in agreement with our list of requirements for a final theory.

Page 149

We have not yet deduced the complete standard model: we still need to show which types of particles exist, which properties they have and what couplings they produce. However, we have found that the strand model explains all the mathematical structures from the millennium list that occur in quantum field theory and in the standard model of particle physics. In fact, the strand explanation for the origin of the gauge interactions allows us to make several definite predictions.

Page 164

Predicting the number of interactions in nature

Already in 1926, Kurt Reidemeister proved an important theorem about possible deformations of knots or tangles that lead to changes of crossings. When tangles are described with two-dimensional diagrams, all possible deformations can be reduced to *exactly three* moves, nowadays called after him. In the strand model, the two-dimensional tangle diagram describes what an observer *sees* about a physical system. Together with the equivalence of interactions as crossing-changing deformations, Reidemeister's theorem thus proves that there are *only three gauge interactions* in nature. In particular, there is no fifth force. Searches for additional gauge interactions are predicted to fail. And indeed, they have all failed up to now.

Ref. 182

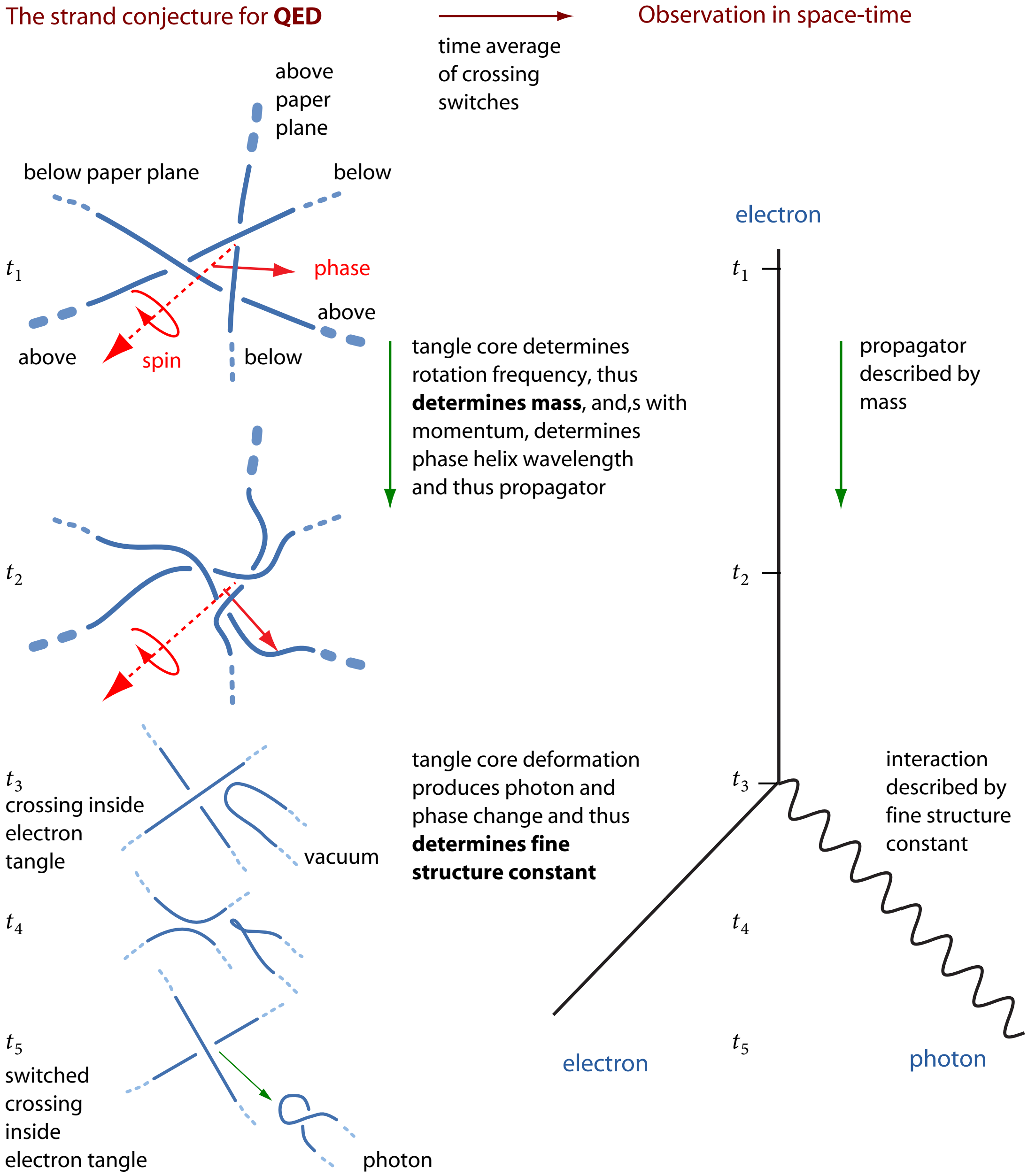


FIGURE 74 QED in one picture: In the strand model, the electron mass and the fine structure constant are determined by the tangle – only the simplest family member is shown here – and its shape change under fluctuations.

Unification of interactions

Ref. 142 We can also state that there is only *one* Reidemeister move. This becomes especially clear if we explore the three-dimensional shape of knots instead of their two-dimensional diagrams: all three Reidemeister moves can be deduced from the *same* deformation of a single strand. Only the projection on a two-dimensional diagram creates the distinction between the three moves. In the terms of the strand model, this means that all gauge interactions are in fact aspects of only one basic process, a fluctuation of strand shape, and that the three gauge interactions are only distinguished by their projections. In this way, the three gauge interactions are thus *unified* by the strand model.

The plane of projection used in a strand diagram defines a mapping from strand fluctuations to Reidemeister moves. The projection plane is defined by the observer, i.e., by the frame of reference. Depending on the projection plane, a general deformation is mapped into different Reidemeister moves. At first sight, the nature of an interaction – whether electromagnetic, strong or weak – seems to depend on the observer. In nature, however, this is not the case. But this contradiction is only apparent. In the strand model, the nature of interaction of a particle results from the type of asymmetry of its tangle core. Certain strand deformations do not lead to interactions, because their effects are suppressed by the averaging of short-time fluctuations underlying every observation. In other words, the averaging process at the basis of observations also ensures that interactions are effectively observer-independent at low energy.

Page 385 In short, the strand model provides a natural *unification* of the interactions. And this unification of the interactions differs completely from any past proposal. The final test, of course, can only be provided by experiment.

No divergences

The strand model implies that there are *no* divergences in the quantum description of nature. This lack of divergence occurs because all measurement values appear after strand effects have been averaged out. As mentioned above, strand effects on space-time Page 153 disappear through 'shivering' and strand effects on particles disappear through wavefunctions.

In summary, in the strand model, no interaction implies or contains divergences: neither gravity nor the gauge interactions. There are neither ultraviolet nor infrared divergences. The strand model avoids divergences, infinities and singularities of any kind from its very start.

Grand unification, supersymmetry and other dimensions

The three gauge interactions are due to the three Reidemeister moves. Therefore, the strand model asserts that there is *no* single gauge group for all interactions. In short, the strand model asserts that there is *no* so-called *grand unification*. The absence of grand unification implies the absence of large proton decay rates, the absence of additional, still undiscovered gauge bosons, the absence of neutron–antineutron oscillations, and the absence of sizeable electric dipole moments in elementary particles. All these searches Page 396 are ongoing at present; the strand model predicts that they yield *null results*.

Supersymmetry and approaches based on it assume gauge group unification. However, as just explained, the strand model predicts that there is no supersymmetry and therefore no supergravity. The strand model also predicts the absence of all conjectured 'superparticles'. In 2016 and again in 2017, the numerous experiments at CERN confirmed the prediction: there is no sign of supersymmetry in nature.

Reidemeister moves are confined to three spatial dimensions. Indeed, the strand

model is based on exactly three spatial dimensions. It predicts that there are no other, undetected dimensions of space. The strand model also predicts the absence of non-
Page 348
commutative space-time, even though, with some imagination, strands can be seen as remotely related to that approach. Finally, the strand model predicts the lack of different vacua: the vacuum is unique.

In short, the strand model differs both experimentally and theoretically from the unification proposals made in the twentieth century. In particular, the strand model predicts the *absence* of additional symmetries, of additional energy scales, and of additional space-time properties at high energy. The strand model predicts that unification is not achieved by searching for higher symmetries, nor for higher dimensions, nor for concepts that contain both. This lack of complex mathematical or symmetry concepts in nature is disappointing; the hopes and search activities in the last fifty years are predicted to have been misguided. In other words, the predictions of the strand model are unpop-
Page 149
ular. However, these predictions agree with our list of requirements for a final theory;
Page 412
and so far, all these predictions agree with experiment.

No new observable gravity effects in particle physics

Page 8
In the 'cube' structure of physics shown in Figure 1, the transition from the final, unified description to quantum field theory occurs by neglecting gravity, i.e., by assuming flat space-time. The same transition occurs in the strand model, where neglecting gravity in addition requires neglecting the strand diameter. In this way, the gravitational constant G disappears completely from the description of nature.

We can summarize our findings on quantum field theory also in the following way:

▷ The strand model predicts that particle masses are the only observable effect of gravity in quantum physics and in particle physics.

Page 311
This result will be complemented below by a second, equally restrictive result that limits the observable quantum effects in the study of gravity. In short, the strand model keeps particle physics and general relativity almost completely separated from each other. This is a consequence of the different effects produced by tail deformations and by core deformations. And again, the prediction of a lack of additional gravitational effects in particle physics agrees with all experiments so far.

The status of our quest

In this chapter, we have deduced that strands predict exactly three interactions. Interactions are deformations of tangle cores and just three classes of such core deformations exist. The three classes of deformations are given by the three Reidemeister moves. Because of the properties of the Reidemeister moves, the three interactions are described by a U(1), a broken SU(2) and a SU(3) gauge symmetry, respectively.

Strands also show that the three interactions are renormalizable, relativistically invariant, and that they follow the least action principle. Strands thus imply the three interaction Lagrangians of the standard model of particle physics. In addition, strands predict the absence of other interactions, symmetries and space-time structures.

Page 164
If we look at the millennium list of open issues in fundamental physics, we have now

solved all issues concerning the mathematical structures that appear in quantum field theory and in the standard model of particle physics.

▷ All mathematical structures found in quantum physics result from the fundamental principle of the strand model.

Equivalently, extension contains all quantum effects. This is an intriguing result that induces us to continue our exploration. Only two groups of issues are still unexplained: the theory of general relativity and the spectrum of elementary particles. We proceed in this order.

Chapter 10

GENERAL RELATIVITY DEDUCED FROM STRANDS

General relativity describes the deformations of the vacuum. In everyday life, gravitation is the only such effect that we observe. But on astronomical scale, gravity shows more phenomena: vacuum can deflect light, producing gravitational lenses, can wobble, giving gravitational waves, and can accelerate, yielding the darkness of the sky and the fascinating black holes. All these observations require general relativity for their description. Therefore, general relativity must be part of any unified description of nature.

In the following, we explain the existence of gravity as a consequence of strands. Then we deduce the field equations of general relativity, the entropy of black holes and relativistic cosmology from the strand model. We also predict the outcome of many quantum gravity experiments. Finally, we deduce the consequences of strands for cosmology. We include several experimental predictions. Of all Planck-scale models of space or space-time, strands seem to be the simplest one that provides these deductions.

Flat space, special relativity and its limitations

Page 209

We have seen above that any observer automatically introduces a 3+1-dimensional *background* space-time. We have also seen that in the case of quantum theory, *physical* space-time, the space-time that is formed by the fluctuations of the vacuum strands, is naturally 3+1-dimensional and flat. In the absence of gravity, physical space and background space coincide.

Page 213

Using strands, we have deduced:

▷ c is the invariant limit for all energy speeds.

This limit is achieved only by free massless particles, such as photons. Strands also showed us that massive particles move more slowly than light. In short, strands reproduce special relativity.

The strand model thus predicts that *pure* special relativity is correct for all situations and all energies in which gravity and quantum theory play no role. The strand model also predicts that when gravity or quantum effects do play a role, general relativity or quantum theory *must* be taken into account. This means that there is no domain of nature in which intermediate descriptions are valid.

It is sometimes suggested that the invariant Planck energy limit for elementary particles might lead to a 'doubly special relativity' that deviates from special relativity

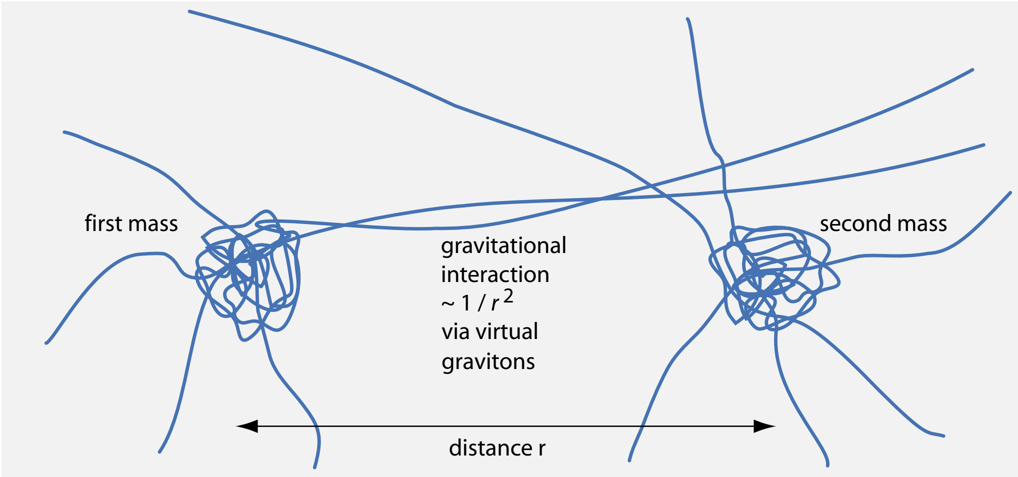


FIGURE 75 Gravitational attraction as result of twisted tail pairs – or twisted tether pairs.

Ref. 85 at high particle energy. However, this suggestion is based on two assumptions: that at Planck energy *point masses* are a viable approximation to particles, and that at Planck energy *vacuum and matter differ*. In the strand model, and in nature, both assumptions are incorrect. Nature, as general relativity shows, does not allow the existence of point masses: the densest objects in nature are black holes, and these are not point-like for any mass value.

In addition, quantum theory implies the fuzziness of matter and space. As a result, Page 65 near Planck energy, matter and vacuum *cannot* be distinguished. Put simply, no system near Planck energy can be described without general relativity and without quantum Ref. 188 gravity. In short, the strand model predicts that the approach of 'doubly special relativ- Page 8 ity' cannot be correct. Also Figure 1 makes this point: there is no description of nature besides the usual ones.

To sum up, the strand model reproduces special relativity when masses are approximated as point-like in flat space. But at the same time, the strand model states that a negligibly small, light and localizable mass cannot exist – neither in flat nor in curved space. This matches observations.

Classical gravitation

In nature, at low speeds and in the flat space limit, gravitation is observed to lead to an acceleration a of test masses that changes as the inverse square distance from the gravitating mass;

$$a = G\frac{M}{R^2} . \tag{169}$$

This acceleration is called *universal gravitation* or *classical gravitation*. It is an excellent approximation for the solar system and for many star systems throughout the universe.

In the strand model, every space-time effect, including gravitation, is due to the behaviour of tangle tails. In the strand model, every mass, i.e., every system of tangles, is connected to the border of space by tails. The nearer a mass is to a second mass, the more frequently the twisted tails of one mass affect the other mass. Figure 75 illustrates the situation. The strand model states:

▷ ***Gravitation*** is due to the fluctuations of tail crossings: gravity is due to twisted tail pairs.

Around a mass, the tail twists fluctuate; averaged of time, the fluctuations lead to a crossing switch density around every mass. The resulting potential energy – where energy is action per time and thus given by the number of crossing switches per time – changes like
Challenge 179 e the inverse distance from the central mass. This is the reason for the $1/r$-dependence of the gravitational potential and the $1/r^2$-dependence of gravitational acceleration. (This applies to all those cases where spatial curvature is negligible.) In simple words, in the strand model, the inverse square dependence of gravitational acceleration is due to the three-dimensionality of space combined with the one-dimensionality of strands.

The strand model also shows that masses and energies are always positive: every tangle contains curved strands. The model also shows qualitatively that larger masses produce stronger attraction: larger masses contain more particles and thus produce more
Page 358 crossing switches. We will show below that the number density of crossing switches for each particle is indeed determined by the mass.

In the strand model, twisted tail pairs are (virtual) gravitons. Strands thus reproduce the idea that gravity is due to the exchange of (virtual) gravitons. Indeed, the strand model of the graviton, illustrated below in Figure 79, provides a consistent model that fulfils all requirements: it has the correct spin and quantum numbers, it fits with the idea of curvature defect, and it couples to masses producing universal gravity.

In the strand model, crossing switches are not only related to action and energy; they are also related to entropy. A slightly different – but equivalent – view on gravitation therefore appears when we put the stress on the entropic aspect.

Deducing universal gravitation from black hole properties

Black holes have entropy; this implies universal gravitation. There are at least two ways to explain this connection.

Ref. 189 An especially concise explanation was recently given by Erik Verlinde. In this view, *gravity appears because any mass M generates an effective vacuum temperature around it.* A gravitating mass M attracts test masses because during the *fall* of a test mass, the total entropy *decreases*. It is not hard to describe these ideas quantitatively.

Given a spherical surface A enclosing a gravitating mass M at its centre, the acceleration a of a test mass located somewhere on the surface is given by the local vacuum temperature T:

$$a = T\,\frac{2\pi\, kc}{\hbar}\ , \qquad (170)$$

where k is the Boltzmann constant. This relation is called the *Fulling–Davies–Unruh effect* and relates vacuum temperature and local acceleration. Thus, an inertial or a freely falling mass (or observer) measures a vanishing vacuum temperature.

In the strand model, the vacuum temperature at the surface of the enclosing sphere is given by the crossing switches induced by the tails starting at the mass. We can determine the vacuum temperature by dividing the energy E contained inside the sphere by *twice* the *maximum* possible entropy S for that sphere. This maximum value is the entropy that the sphere would have if it were a black hole horizon; it can be calculated by the strand model, as we will see shortly.

Page 289

The temperature T is thus given by the expression

$$T = \frac{E}{2S} = \frac{M}{A}\frac{2G\hbar}{kc} . \qquad (171)$$

The factor 2 needs explanation; it derives from the dependence of entropy on the square of the radius. We do not discuss the details here.

Neglecting spatial curvature, we can set $A = 4\pi R^2$; this gives a temperature at the enclosing sphere given by

$$T = \frac{M}{R^2}\frac{G\hbar}{2\pi\, kc} . \qquad (172)$$

Page 283

Inserting this expression into the expression (170) for the Fulling–Davies–Unruh acceleration a, we get

$$a = G\frac{M}{R^2} . \qquad (173)$$

This is universal gravitation, as discovered by Robert Hooke and popularized by Isaac Newton. Since spatial curvature was neglected, and the central mass was assumed at rest, this expression is only valid for large distances and small speeds. We have thus deduced universal gravity from the effects of gravitating masses on vacuum temperature. Below, we show that in the relativistic case this sequence of arguments – which was given by Jacobson fifteen years before Verlinde – leads to the field equations of general relativity.

Page 295

An alternative deduction of universal gravitation from black hole entropy is the following. The gravitational force F on a test mass m is given by the vacuum temperature T created by the central mass M and by the change of entropy S per length that is induced by the motion of the test mass:

$$F = T\frac{\mathrm{d}S}{\mathrm{d}x} . \qquad (174)$$

The change of entropy $\mathrm{d}S/\mathrm{d}x$ when a test mass m moves by a distance x can be determined from the strand model in a simple manner. When the test mass m moves by a (reduced) Compton wavelength, in the strand model, the mass has rotated by a full turn: the entropy change is thus $2\pi k$ per (reduced) Compton wavelength. Thus we have

$$\frac{\mathrm{d}S}{\mathrm{d}x} = m\frac{2\pi\, kc}{\hbar} . \qquad (175)$$

Using the temperature T found in expression (172), we get an expression for the gravitational force given by

$$F = G\frac{Mm}{R^2} . \tag{176}$$

This is universal gravitation again.

We have thus deduced universal gravitation from the entropy and the vacuum temperature generated by gravitating masses. We note that the temperature and entropy of black holes are limit values. We can thus state that universal gravitation is a consequence of nature's limit values. Page 36

Summary on universal gravitation from strands

Universal gravitation is due to the temperature and entropy of the (curved) vacuum around masses. The limit case is the temperature and entropy of black holes. In the strand model, these temperature and entropy values are a consequence of the underlying strand crossing switches; we will show this shortly. Page 289

More precisely, gravitation is due to twisted tail pairs. In the strand model, universal gravitation thus appears as an effect of the crossing switches induced by masses. In fact, we have several explanations of universal $1/r^2$ gravitation using strands. We have deduced universal gravitation from the energy of strands, from the temperature of strands and from the entropy of strands around a mass. We have also have deduced universal gravitation from the maximum force, which strands fulfil as well. In short, strands explain the origin of universal gravitation in several consistent ways.

Incidentally, modelling mass as a source for strand crossing switches – twisted strand pairs – is remotely reminiscent of Georges-Louis Lesage's eighteenth-century model of gravitation. Lesage proposed that gravity appears because many tiny, usually unnoticed corpuscules push masses together. In fact, as we will see shortly, there is a certain similarity between these assumed tiny corpuscules and virtual gravitons. And interestingly, all criticisms of Lesage's model then cease to hold. First, there is no deceleration of free masses in inertial motion, thanks to the built-in special-relativistic invariance. Secondly, there is no heating of masses, because the entangled tails represent virtual gravitons that scatter elastically. Thirdly, and most of all, by replacing the *corpuscules ultra-mondains* of Lesage by virtual gravitons – and finally by strands – we can predict an additional effect of gravity that is not described by the inverse square dependence: space-time curvature. Ref. 190 Vol. I, page 218

Curved space

In nature, observation shows that physical space is not flat around masses, i.e., in the presence of gravity. Near mass and energy, physical space is *curved*. Observations also show that curved space-time remains 3+1-dimensional. The observation of this type of curvature was predicted long before it was measured, because curvature follows unambiguously when the observer-invariance of the speed of light c and the observer-invariance of the gravitational constant G are combined.

We continue directly with the strand model of spatial curvature and show that all observations are reproduced.

▷ ***Curvature*** (of physical space-time) is due to simple, unknotted and weakly

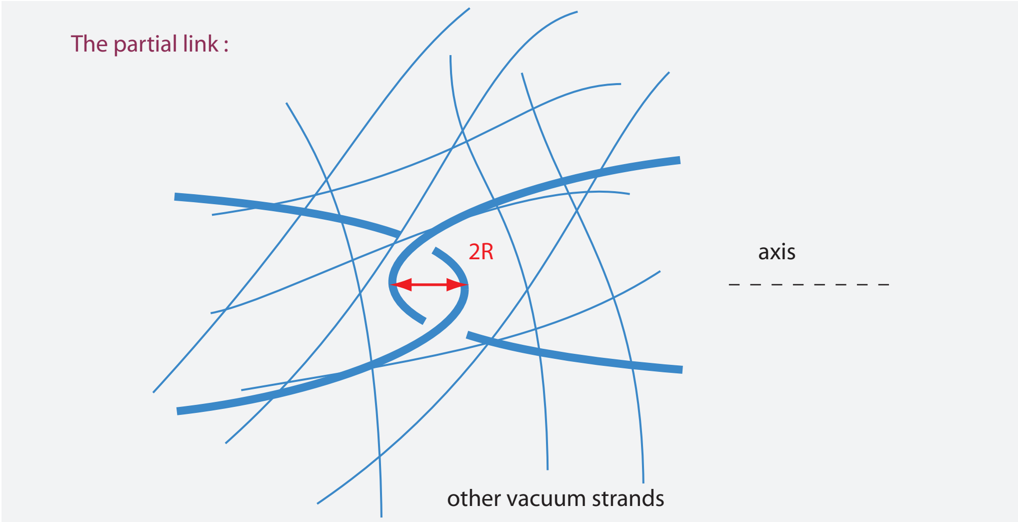


FIGURE 76 A schematic model of the fundamental defect, and thus the fundamental type of curvature: the *partial link*.

localized defects in the tangle of strands that make up the vacuum. An example is shown in Figure 76.

▷ In the case of curvature, *physical* space-time, which is due to averaged strand crossing switches, *differs* from flat *background* space-time, which usually corresponds to the tangent or to the asymptotic space-time. In Figure 76, the grey background colour can be taken as visualization of the background space.

▷ ***Mass*** is a localized defect in space and is due to tangled strands. Thus mass curves space around it.

▷ ***Energy*** in a volume is the number of crossing switches per unit time. As a result, mass is equivalent to energy. As a second result, energy also curves space.

▷ ***Gravitation*** is the space-time curvature originating from compact regions with mass or energy.

These natural definitions show that curvature is due to strand configurations. In particular, curvature is built of unknotted – i.e., massless – *defects*. The massless defects leading to curvature are usually dynamic: they evolve and change. Such curvature defects – virtual gravitons – originate at regions containing matter or energy. In fact, the curvature of Page 358 space around masses is a natural result of fluctuations of the strands that make up matter tangles.

We note that curved space, being a time average, is *continuous* and *unique*. Vacuum or curved space, more precisely, curved physical space, thus differs from background space, which is flat (and drawn in grey in the figures).

Incidentally, the distinction between physical and background space also avoids Einstein's hole argument; in fact, the distinction allows discussing it clearly, as only physical

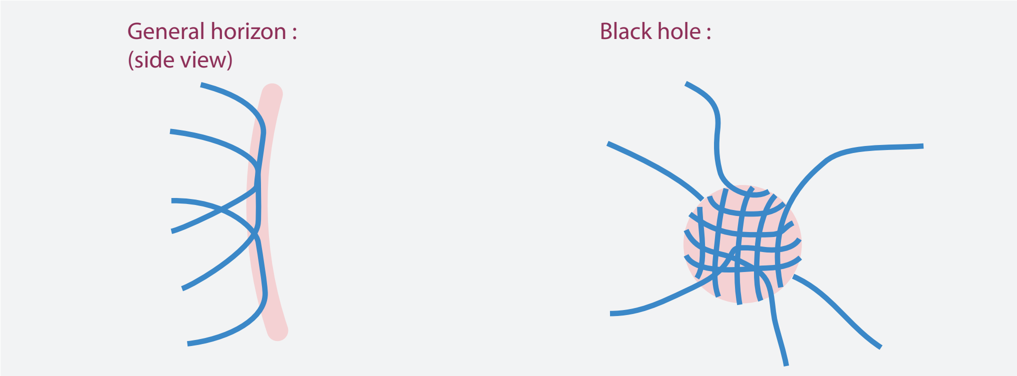


FIGURE 77 A schematic model of a general and a spherical horizon as tight weaves, as pictured by a distant observer. In the strand model there is *nothing*, no strands and thus no space, behind a horizon.

Vol. II, page 284
space describes nature.

The structure of horizons and black holes

In general relativity, another concept plays a fundamental role. In the strand model we have:

▷ A ***horizon*** is a tight, one-sided weave of strands.

Therefore, there are no strands behind the horizon. This implies that behind a horizon, there is no matter, no light, no space and no time – just *nothing*. Indeed, this is the experience of any observer about a horizon. A horizon is thus a structure that limits physical space. It does *not* limit background space.

One particular type of horizon is well-known:

▷ A ***black hole*** is a tight, one-sided and *closed* weave of strands.

In principle, closed horizons can have any shape. The simplest case is the spherical, non-rotating horizon, which defines the *Schwarzschild black hole*. It is illustrated on the right-hand side of Figure 77.

If an observer is located outside a spherical horizon, the strand model states that there is nothing *inside* the horizon: no matter, no light and no vacuum. The strand model thus provides a simple and drastic view of black hole horizons. Figure 77 also illustrates that the concept of radius (or size) of a black hole has to be approached with the (well-known) care. In general, the size of a structure made of strands is the number of crossings encountered when travelling through it. However, an observer cannot travel *through* a black hole: there are no strands inside, thus there is no vacuum there! The size of a black hole must therefore be defined indirectly. The simplest way is to take the square root of the area, divided by 4π, as the radius. Thus the strand model, like general relativity, requires that the size of a compact horizon be defined by travelling *around* it.

We note that the strand model also provides an intuitive explanation for the differences between a rotating and a non-rotating black hole.

Is there something behind a horizon?

A drawing of a horizon weave, such as the one of Figure 77, clearly points out the difference between the background space and the physical space. The *background space* is the space we need for thinking, and is the space in which the drawing is set. The *physical space* is the one that appears as a consequence of the averaging of the strand crossings. Physical, curved space exists only on the observer side – usually outside – of the horizon. The physical space around a black hole is curved; it agrees with the background space only at infinite distance from the horizon. Inside the horizon, there is background space, but no physical space. In short, the strand model implies that – for an observer at spatial infinity – there is *nothing*, not even a singularity, inside a black hole horizon.

▷ There is no physical space, no matter and no singularity inside a horizon.

Horizons are observer-dependent. Both the existence and the shape of a horizon depends on the observer. As we will see, this happens in precisely the same way as in usual general relativity. In the strand model, there is no contradiction between the one observer at spatial infinity who says that there is *nothing* behind a horizon, not even physical space, and another, falling observer, who does not observe a horizon and thus states that there is *something* there. In the strand model, the two statements naturally transform into each other under change of viewpoint. Indeed, the transformation between the two viewpoints contains a deformation of the involved strands.

We note that the equivalence of viewpoints and the statement that there is nothing behind a horizon is based on the combination of general relativity and quantum theory. If we would continue thinking that space and time is a manifold of points – thus disregarding quantum theory – these statements would *not* follow.

In summary, one-sided tight weaves are a *natural* definition of horizons.

Energy of black hole horizons

The strand model allows us to calculate the energy content of a closed horizon. Energy is action per unit time. In the strand model, the energy of a non-rotating spherical horizon is thus given by the number N_{cs} of crossing switches per time unit. In a tight weave, crossing switches cannot happen in parallel, but have to happen sequentially. As a result, a crossing switch 'propagates' to the neighbouring Planck area on the surface. Since the horizon weave is tight and the propagation speed is one crossing per crossing switch time, this happens at the speed of light. In the time T that light takes to circumnavigate the spherical horizon, all crossings switch. We thus have:

Challenge 180 e

$$E = \frac{N_{cs}}{T} = \frac{4\pi R^2}{2\pi R}\,\frac{c^4}{4G} = R\,\frac{c^4}{2G}\,. \qquad (177)$$

Strands thus imply the well-known relation between energy (or mass) and radius of Schwarzschild black holes.

How do the crossing switches occur at a horizon of a black hole? This interesting

Challenge 181 e

puzzle is left to the reader.

The tight-weave model of horizons also illustrates and confirms both the *hoop conjecture* and the *Penrose conjecture*. For a given mass, because of the minimum size of crossings, a spherical horizon has the smallest possible diameter, compared to other possible shapes. The strand model naturally implies that, for a given mass, spherical black holes indeed are the densest objects in nature.

The nature of black holes

The strand model naturally implies the *no-hair theorem*. Since all strands are the same, independently of the type of matter that formed or fell into the horizon, a black hole has no characteristics other than mass, angular momentum and charge. Here we used a result from the next chapter, when it will become clear that all elementary particles are indeed made of the same featureless strands. Taking that result as given, we deduce that flavour quantum numbers and particle number do not make sense for black holes. We also deduce that weak and strong charge are not defined for black holes. Strands explain naturally why neutral black holes made of antimatter and neutral black holes made of matter do not differ, if their masses and angular momenta are the same. In short, the strand model of nature implies the no-hair theorem: *strands, not hairs*.

Horizons and black holes are borderline systems between space and matter. This borderline property must be fulfilled by every final theory. The strand model fulfils this requirement: in the strand model, black holes can either be described as curved space or as tightly packed particles in permanent free fall.

Entropy of vacuum and matter

Both vacuum and matter are made of fluctuating strands. We note directly:

▷ The flat and infinite vacuum has *vanishing* entropy, because the number of crossing switches is zero on average.

At the same time,

▷ Curved space and horizons have *non-vanishing* entropy.

The entropy of vacuum and of horizons differs from that of matter. In the *absence of gravity*, the number of microstates of matter is determined – as in usual thermodynamics (thermostatics) – by the behaviour of tangle *cores*.

In *strong gravity*, when the distinction between matter and vacuum is not so clear-cut, the number of microstates is determined by the possible crossing switches of the strands. In strong gravity, only *tails* play a role. This becomes clear when we calculate the entropy of black holes.

Entropy of black holes deduced from the strand model

Despite the tight weaving, the strands making up a horizon are fluctuating and moving: the weave shape fluctuates and crossing switch all the time. This fluctuating motion is

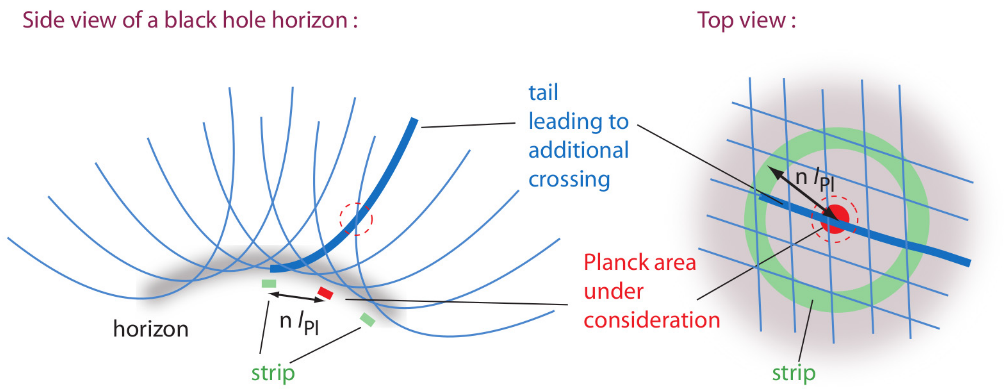


FIGURE 78 The entropy of black holes results from the number of possible crossing states above a Planck area.

the reason why horizons – in particular those of black holes – have entropy.

The weave model of a horizon, illustrated in detail in Figure 78, allows us to calculate the corresponding entropy. Since the horizon is a tight weave, there is a crossing on each Planck area. To a first approximation, on each (corrected) Planck area of the horizon, the strands can cross in *two* different ways. The fundamental principle of the strand model thus yields two microstates per Planck area. The number N of Planck areas is given by $N^2 = Ac^3/4G\hbar$. The resulting number of black hole microstates is 2^{N^2}. The entropy is given by the natural logarithm of the number of the possible microstates times k. This approximation gives an entropy of a horizon of

$$S = A\,\frac{kc^3}{4G\hbar}\ln 2\,. \tag{178}$$

This result is the well-known first approximation of black hole entropy: one bit per corrected Planck area. In the strand model, the proportionality of entropy and area is thus a direct consequence of the *extension* of the strands. This proportionality is also well known from studies of quantum gravity and of strings. In those approaches however, the relation between the area proportionality and extension is less obvious.

Ref. 191

For Schwarzschild black holes, the entropy value of expression (178) is *not* correct. In the strand model, this incorrect value is explained as a consequence of neglecting the effects of the strand *tails*. Indeed, additional contributions to the entropy appear at a *finite distance* from the horizon, due to the crossing of the tails on their way to the border of space, as shown in Figure 78. The actual entropy will thus be larger than the first approximation, but still be proportional to the area A.

The correct proportionality factor between the area and the entropy of a black hole results when the strand tails are taken into account. (The correction factor is called the *Barbero–Immirzi parameter* in the research literature on quantum gravity.) The calculation is simplest for Schwarzschild black holes. By construction, a black hole with macroscopic radius R, being a tight weave, has R/l_{Pl} tails. For each given Planck area, there are, apart from the basic, or lowest crossing, additional crossings 'above it', along the radial direction, as shown in Figure 78. These additional crossings are due to the tails from neighbouring and distant Planck areas.

Taking into effect all strand tails allows us to calculate the average number of crossings *above* a given Planck area. The main point is to perform this calculation for all those tails that start in a circular strip of Planck width centred around the Planck area under consideration. We then add the probabilities for all possible circular strips. One such circular strip is drawn in Figure 78.

The definition of horizons as tight weaves implies that a horizon with N^2 Planck areas is made of N strands. This means that for each circular strip of radius nl_{Pl}, there is only *one* strand that starts there and reaches spatial infinity as a tail. For this tail, the average

Challenge 182 e probability p that it crosses above the central Planck area under consideration is

$$p = \frac{1}{n!} \ . \tag{179}$$

Summing over all strips, i.e., over all values n, we get a total of $\sum_{n=0}^{\infty} 1/n! = \mathrm{e} = 2.71828...$ microstates *on* and *above* the central Planck area under consideration. Thus the number e replaces the number 2 of the first approximation: the number of horizon microstates of a Schwarzschild black hole is not 2^{N^2}, but e^{N^2}. As a consequence, the entropy of a macroscopic Schwarzschild horizon becomes

$$S = A\,\frac{kc^3}{4G\hbar} \ . \tag{180}$$

This is the Bekenstein–Hawking expression for the entropy of Schwarzschild black holes. The strand model thus reproduces this well-known result. With this explanation of the difference between 2 and e = 2.71828..., the strand model confirms an old idea:

▷ The entropy of a black hole is located *at and near* the horizon.

The above calculation, however, counts some states more than once. Topologically identical spherical horizons can differ in the direction of their north pole and in their state of rotation around the north–south axis. If a spherical horizon is made of N strands, it has N^2 possible physical orientations for the north pole and N possible angular orientations around the north–south axis. The actual number of microstates is thus e^{N^2}/N^3. Using the relation between N^2 and the surface area A, namely $A = N^2 4G\hbar/c^3$, we get the final result

$$S = A\,\frac{kc^3}{4G\hbar} - \frac{3k}{2}\ln\frac{A\,c^3}{4G\hbar} \ . \tag{181}$$

The strand model thus makes a specific prediction for the logarithmic correction of the entropy of a Schwarzschild black hole. This final prediction of the strand model agrees with many (but not all) calculations using superstrings or other quantum gravity

Ref. 192 approaches.

In summary, the entropy value (180), respectively (181), of black holes is due to the *extension* of the fundamental entities in the strand model and to the *three dimensions* of space. If either of these properties were not fulfilled, the entropy of black holes would not result. This is not a surprise; also our deduction of quantum theory was based on the same two properties. In short: like every quantum effect, also the entropy of black holes is a result of extension and three-dimensionality. Only a three-dimensional description of nature agrees with observation.

Temperature, radiation and evaporation of black holes

The strands that make up a horizon fluctuate in shape. Since every horizon contains energy, the shape fluctuations imply energy fluctuations. In other words, horizons are predicted to have a *temperature*. The value of the temperature can be deduced from the strand model by noting that the characteristic size of the fluctuations for a spherical horizon is the radius R of the horizon. Therefore we have

$$kT = \frac{\hbar c}{2\pi R} \ . \tag{182}$$

Using the definition of *surface gravity* as $a = c^2/R$, we get

$$T = \frac{\hbar a}{2\pi k c} \ . \tag{183}$$

Ref. 57, Ref. 58 The strand model predicts that horizons have a temperature proportional to their surface gravity. This result has been known since 1973.

All hot bodies radiate. The strand model thus predicts that Schwarzschild black holes *radiate* thermal radiation of the horizon temperature, with power and wavelength

$$P = 2\pi\hbar c^2/R^2 \quad , \quad \lambda \approx R \ . \tag{184}$$

This confirms a well-known consequence of the temperature of black holes.

Like all thermal systems, horizons follow thermodynamics. In the strand model, black hole radiation and evaporation occur by reduction of the number of strands that make up the horizon. The strand model thus predicts that black holes *evaporate completely*, until only elementary particles are left over. In particular, the strand model implies that in black hole radiation, there is *no* information loss.

In short, strands reproduce all aspects of black hole evaporation. The strand model also shows that there is no information loss in this process.

Black hole limits

In many ways, black holes are *extreme* physical systems. Not only are black holes the limit systems of general relativity; black holes also realize various other limits. As such, black holes resemble light, which realizes the speed limit. We now explore some of these limits.

For a general physical system, not necessarily bound by a horizon, the definitions of energy and entropy with strands allow some interesting conclusions. The entropy of a system is the result of the number of crossing possibilities. The energy of a system is the number of crossing changes per unit time. A large entropy is thus only possible if a system shows many crossing changes per time. Since the typical system time is given by the circumference of the system, the entropy of a physical system is therefore limited:

$$S \leqslant ER\, 2\pi k/\hbar c\ . \tag{185}$$

This relation is known as *Bekenstein's entropy bound*; the precise definitions of the quantities in the bound need some care, as Don Page explains. The bound thus also follows from the strand model. Strands imply that the equality is realized only for black holes.

Ref. 193

In the strand model, horizons are tight, one-sided weaves. For example, this implies that any tangle that encounters a horizon is essentially flat. Because of tangle flatness and the extension of the tails, at most one Planck mass can cross a horizon during a Planck time. This yields the mass rate limit

$$\mathrm{d}m/\mathrm{d}t \leqslant c^3/4G \tag{186}$$

that is valid in general relativity and in nature.

Black holes can rotate. The strand model states that there is a highest angular frequency possible; it appears when the equator of the black hole rotates with the speed of light. As a result, the angular momentum J of a black hole is limited by

$$J < 2GM^2/c\ . \tag{187}$$

Ref. 194

This limit is well known from general relativity.

The electric charge of a black hole is also limited. The force limit in nature implies that the electrical forces between two charged black holes must be lower than their gravitational interaction. This means that

$$\frac{Q^2}{4\pi\varepsilon_0 r^2} \leqslant \frac{GM^2}{r^2}\ , \tag{188}$$

or

$$Q^2 \leqslant 4\pi\varepsilon_0 GM^2\ . \tag{189}$$

This is the well-known charge limit for (static) black holes given by the Reissner–Nordström metric. The maximum charge of a black hole is proportional to its radius. It follows directly from the maximum force principle.

To explain the charge limit, we deduce that the *extremal* charge surface density Q/A of a black hole is proportional to $1/R$. The higher the horizon curvature, the more charge per Planck area is possible. In the strand model, a horizon is a tight weave of strands. We are thus led to conjecture that at Planck scale, electric charge is related to and limited by strand curvature. We will explore this connection in more detail below.

Page 382

The strand model limits energy density to the Planck energy per Planck volume, or to the value $c^7/(16G^2\hbar)$. This limit implies a lower size limit for black holes, particles and any localized system. Therefore, the strand model does not allow singularities, be they dressed or naked. And indeed, no singularity has ever been observed.

In summary, the strand model reproduces the known limit properties of horizons. And all these results are independent of the precise fluctuation details of the strands.

Curvature around black holes

The tails of a black hole extend up to the border of space; the density of tails is highest at the horizon. A black hole is therefore surrounded by partial links at any *finite* distance from the horizon. In other words, the space around a black hole is *curved*. The value of the space-time curvature increases as one approaches the horizon, because of the way in which the partial links hinder each other in their motion. The nearer they are to the horizon, the more they hinder each other. The curvature that appears is proportional to the density of partial links and to their average strand curvature.

At the horizon, the curvature radius is the horizon radius R. By construction, the number of tails departing from a non-rotating black hole is proportional to R. The spatial curvature is given by the average crossing density gradient. Hence at a radial distance r from a static black hole, the spatial curvature K is

$$K \sim \frac{R}{r^3} . \tag{190}$$

So at the horizon itself, the curvature K is (of the order of) the inverse square of the horizon radius; further away, it decreases rapidly, with the third power of the distance. This result is a well-known property of the Schwarzschild solution and is due to the extension of the strands. The rapid decay with radius is the reason why in everyday situations there is no noticeable curvature of space-time. In short, strands allow us to deduce the correct curvature of space-time around black holes and spherical masses.

The shape of non-rotating black holes

The strand model also explains and visualizes the importance of spherical horizons in nature. First of all, strands illustrate the non-existence of (uncharged) one-dimensional or toroidal horizons in 3 + 1 space-time dimensions. Such configurations are unstable, in particular against transverse shear and rearrangement of the strands.

The strand model also implies that non-rotating, closed horizons are spherical. Obviously, spheres are the bodies with the smallest surface for a given volume. The minimum horizon surface appears because the strands, through their fluctuations, effectively 'pull' on each Planck area of the horizon. As a result, all non-rotating macroscopic horizons will evolve to the spherical situation in a few Planck times. (Deviations from the spherical shape will mainly occur near Planck scales.) With the definition of gravity waves given below, it also becomes clear that strongly deformed, macroscopic and non-spherical horizons are unstable against emission of gravity waves or of other particles. In short,

▷ All non-rotating horizons of non-spherical shape are unstable.

The strand model thus confirms that spherical horizons are favoured and that the most compact bodies with a given mass. The reasoning can be extended to rotating horizons, yielding the well-known shapes.

In summary, strands reproduce all known qualitative and quantitative properties of horizons and of black holes, and thus of general systems with strong gravitational fields. All predictions from strands agree with observations and with other approaches to quantum gravity. These hints already suggest that strands imply the field equations.

The field equations of general relativity

The field equations can be deduced from the fundamental principle in two different, but related ways. Essentially, both derivations repeat the reasoning for universal gravitation given above, but for the relativistic case. The first deduction of the field equations is based on an old argument on the thermodynamics of space-time. Strands show that horizons have three thermodynamic properties:

Page 282
Ref. 22

— an area–entropy relation of $S = A\, kc^3/4G\hbar$,

— a curvature–temperature relation of $T = a\,\hbar/2\pi kc$,

— a relation between heat and entropy of $\delta Q = T\delta S$.

Using these three properties, and using the relation

$$\delta Q = \delta E \;, \tag{191}$$

that is valid *only* in case of horizons, we get the first principle of horizon mechanics

$$\delta E = \frac{c^2}{8\pi G} a\, \delta A \;. \tag{192}$$

From this relation, using the Raychaudhuri equation, we obtain the field equations of general relativity. This deduction was given above.*

Page 33

In other words, the field equations result from *the thermodynamics of strands*. It is worth noting that the result is independent of the details of the fluctuations or of the microscopic model of space, as long as the three thermodynamic properties just given

* Here is the argument in a few lines. The first principle of horizon mechanics can be rewritten, using the energy–momentum tensor T_{ab}, as

$$\int T_{ab} k^a \mathrm{d}\Sigma^b = \frac{c^2}{8\pi G} a\, \delta A$$

where $\mathrm{d}\Sigma^b$ is the general surface element and k is the Killing vector that generates the horizon. The Raychaudhuri equation allows us to rewrite the right-hand side as

$$\int T_{ab} k^a \mathrm{d}\Sigma^b = \frac{c^4}{8\pi G} \int R_{ab} k^a \mathrm{d}\Sigma^b$$

where R_{ab} is the Ricci tensor describing space-time curvature. This equality implies that

$$T_{ab} = \frac{c^4}{8\pi G}(R_{ab} - (R/2 + \Lambda) g_{ab})$$

where Λ is an undetermined constant of integration. These are Einstein's field equations of general relativity. The field equations are valid everywhere and for all times, because a suitable coordinate transformation can put a horizon at any point and at any time. To achieve this, just change to a suitable accelerating frame, as explained in the volume on relativity.

Vol. II, page 98

are valid. In fact, these properties must be fulfilled by any model of space-time; and indeed, several competing models of space claim to fulfil them.

We can use the relation between fluctuations and strands to settle an issue mentioned
Page 188
above, in the section on quantum theory. Strand fluctuations *must* obey the thermodynamic properties to allow us to define space-time. If they obey these properties, then space-time exists and curves according to general relativity.

A second derivation of the field equations of general relativity follows the spirit of the strand model most closely. It is even shorter. Strands imply that all physical quantities are limited by the corresponding Planck limit. These limits are due to the limit to the fundamental principle, in other words, they are due to the packing limit of strands. In
Ref. 19
particular, the fundamental principle limits force by $F \leqslant c^4/4G$ and power by $P \leqslant c^5/4G$.
Page 33
We have already shown above that this limit implies the field equation.

In other words,

▷ Given that black holes and thus horizons are thermodynamic systems, so is curved space.

The reason: both can be transformed into each other. Therefore:

▷ Since black holes have thermodynamic aspects, so has gravity.

And since black holes are built from microscopic degrees of freedom, so is curved space. Or, in simple words:

▷ Space is made of many small entities.

And finally we can state:

▷ Space is made of strands, because strands are the simplest entities that yield black hole entropy.

Strands are the simplest way to incorporate quantum effects into gravitation. If we take into consideration that strands are the only way known so far to incorporate gauge interactions, we can even conclude that strands are the only way known so far to incorporate all quantum effects into gravitation.

In summary, the strand model asserts that the field equations appear as consequences of fluctuations of impenetrable, featureless strands. In particular, the strand model implies and confirms that a horizon and a particle gas at Planck energy do not differ. However, the value of the cosmological constant is *not* predicted from strand thermodynamics.

Equations from no equation

The strand model asserts that the field equations of general relativity are not the result of another, more basic evolution equation, but result directly from the fundamental principle. To say it bluntly, the field equations are deduced from a drawing – the funda-

Page 149 mental principle shown in Figure 10. This strong, almost unbelievable statement is due to a specific property of the field equations and to two properties of the strand model.

First of all, the field equations are, above all, consequences of the thermodynamics of space-time. In the strand model, the thermodynamic properties are deduced as a consequence of the strand fluctuations. This deduction does not require underlying evolution equations; the field equations follow from the statistical behaviour of strands.

The second, essential property of the strand model is its independence from the underlying motion of the strands. In the strand model we obtain the evolution equations of the vacuum – the field equations of general relativity – without deducing them from another equation. We do not need an evolution equation for the strand shape; the deduction of the field equations works for *any* underlying behaviour of strand shapes, as long as the thermodynamic properties of the strand fluctuations are reproduced.

The third and last essential property that allows us to deduce the field equations directly from a graph, and not from another equation, is the relation between the graph and natural physical units. The relation with natural units, in particular with the quantum of action $\hbar$ and the Boltzmann constant k, is fundamental for the success of the strand model.

In summary, the fundamental principle of the strand model contains all the essential properties necessary for deducing the field equations of general relativity. In fact, the discussion so far makes another important point: unique, underlying, more basic evolution equations for the tangle shape *cannot* exist. There are two reasons. First, an underlying equation would itself require a deduction, thus would not be a satisfying solution to unification. Secondly, and more importantly, evolution equations are differential equations; they assume well-behaved, smooth space-time. At Planck scales, this is impossible.

▷ Any principle that allows deducing the field equations cannot itself be an evolution equation.

The Hilbert action of general relativity

We have just shown that the strand model implies the field equations of general relativity.
Page 211 We have also shown above that, in the strand model, the least action principle is a natural property of all motion of strands. Combining these two results, we find that a natural way to describe the motion of space-time is the (extended) *Hilbert action* given by

$$W = \frac{c^4}{16\pi G} \int (R - 2\Lambda)\, \mathrm{d}V \ , \tag{193}$$

where R is the Ricci scalar, $\mathrm{d}V = \sqrt{\det g}\, \mathrm{d}^4 x$ is the invariant 4-volume element of the metric g, and Λ is the cosmological constant, whose value we have not determined yet. As is well known, the description of evolution with the help of an action does not add anything to the field equations; both descriptions are equivalent.

For a curved three-dimensional space, the Ricci scalar R is the average amount, at a given point in space, by which the curvature deviates from the zero value of flat space.
Page 286 In the strand model, this leads to a simple statement, already implied by Figure 76:

FIGURE 79 The graviton in the strand model: a twist in a pair of tethers.

▷ The *Ricci scalar* R is the ratio of additional or missing crossings per spatial volume, compared to flat space.

As usual, the averaging is performed over all spatial orientations. A similar statement can be made for the cosmological constant Λ. In short, we can say: the Hilbert action follows directly from the fundamental principle of the strand model.

Space-time foam

Quantum physics implies that at scales near the Planck length and the Planck time, space-time fluctuates heavily. John Wheeler called the situation *space-time foam*; the term *quantum foam* is also used. In a sense, *quantum gravity* can be defined, if at all, as the description of space-time foam. This reduced view arises because no separate theory of quantum gravity is possible in nature.

Historically, there have been many speculations on the details of space-time foam. Apart from its fluctuations, researchers speculated about the appearance of topology changes – such as microscopic wormholes – about the appearance of additional dimensions of space – between six and twenty-two – or about the appearance of other unusual properties – such as microscopic regions of negative energy, networks or loop structures.

The strand model makes a simple prediction that contradicts most previous speculations:

▷ Space-time foam is made of fluctuating strands.

At everyday scales, the foam is not noticed, because background space and physical space are indistinguishable. At Planck scales, space-time is not fundamentally different from everyday space-time. No unusual topology, no additional dimensions, and no new or unusual properties appear at Planck scales. Above all, the strand model predicts that there are *no* observable effects of space-time foam; for example, 'space-time noise' or 'particle diffusion' do not exist. The strand model of space-time foam is both simple and unspectacular. Ref. 195

Gravitons, gravitational waves and their detection

In the strand model, gravitons can be seen as a special kind of partial links: it appears to be a twisted pair of tethers. An example is shown in Figure 79. As a twisted pair of parallel strands, the graviton returns to itself after rotation by π; it thus behaves like a spin-2 boson, as required.

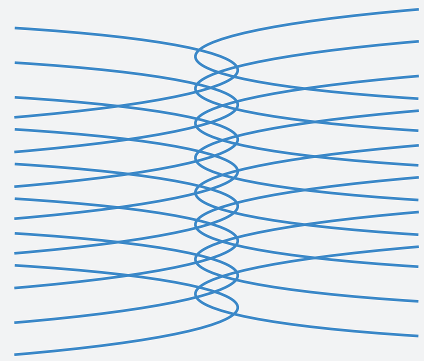


FIGURE 80 A speculative, highly schematic model for a cosmic string, a one-dimensional defect in space-time

Can single gravitons be observed? The strand model implies that the absorption of a single graviton by an elementary particle changes its spin or position. However, such a change *cannot* be distinguished from a quantum fluctuation, because the graviton is predicted to be massless. Furthermore, the strand model predicts that gravitons do not interact with photons, because they have no electric charge. In summary, the strand model predicts:

▷ Single gravitons *cannot* be detected.

The situation changes for gravitational waves. Such waves are coherent superpositions of large numbers of gravitons and are observable classically. In such a case, the argument against the detection of single gravitons does not apply. In short, the strand model predicts that gravitational waves *can* be observed. (This prediction, made by many since 1915 and repeated in this text on the basis of the strand model in 2008, came true in February 2016. The observations also produced the extremely low mass limit of at most $1.2 \times 10^{-22}\,\mathrm{eV}/c^2$ for any possible mass of the photon.)

Challenge 183 e

Ref. 196

Open challenge: Improve the argument for the graviton tangle

The argument that leads to the graviton tangle is rather hand-waving. Can you make the argument more compelling? Could the four tails form a cross and thus span a plane instead of a ribbon?

Challenge 184 ny

Other defects in vacuum

The strand model provides a quantum description of gravitation. The strand model does so by explaining physical space as the average of the crossing switches induced by strand fluctuations among untangled strands. Matter, radiation and horizons are defects in the 'sea' of untangled strands.

So far, we have been concerned with *particles*, i.e., localized, zero-dimensional defects, and with *horizons*, i.e., two-dimensional defects. Now, modelling of the vacuum as a set of untangled strands also suggests the possible existence of *one-dimensional* – equivalent to dislocations and disclinations in solids – of *additional* two-dimensional defects, or of *three-dimensional* defects. Such defects could model cosmic strings, domain walls, wormholes, toroidal black holes, time-like loops and regions of negative energy.

An example of such a possible new defect is illustrated in Figure 80. The illustration can be seen as the image of a one-dimensional defect or as the cross section of a two-dimensional defect. Are such defects stable against fluctuations? The strand model suggests that they are not. These defects are expected to decay into a mixture of gravitons, black holes, matter and radiation particles. However, this issue is still a topic of research, and will not be covered here.

Exploring the stability of wormholes, time-like loops and toroidal black holes leads to similar results. It seems that the strand model should not allow time-like loops of macroscopic size, since any configuration that cannot be embedded locally into three flat spatial dimensions is either a particle or a black hole. Alternatively, macroscopic time-like loops would collapse or decay because of the fluctuations of the strands. In the same way, wormholes or black holes with non-trivial topology should be unstable against more usual strand structures, such as particles or black holes.

We also note the strand model does not allow volume defects (black holes being surface-like defects). The most discussed types of volume defects are macroscopic regions of negative energy. Energy being action per unit time, and action being connected to crossing changes, the model does not allow the construction of negative-energy regions. However, the strand model does allow the construction of regions with lower energy than their environment, as in the Casimir effect, by placing restrictions on the wavelengths of photons.

The strand model thus predicts the absence of additional defects and tangle types. The final and general connection between tangle types and defects is shown (again) in Table 11. The next chapter will give details of the tangles corresponding to each particle.

In summary, the strand model reproduces the results of modern quantum gravity and predicts that the more spectacular defects conjectured in the past – linear defects such as cosmic strings, surface defects such as wormholes, volume defects such as negative-energy regions – *do not appear* in nature.

The gravity of superpositions

What is the gravitational field of a quantum system in a macroscopic superposition? The issue has been raised by many scholars as an important step towards the understanding of how to combine gravitation and quantum theory.

The strand model deflates the importance of the issue. The model shows – or predicts, if one prefers – that the gravitational field of a superposition is the temporal and spatial average of the evolving quantum system, possibly under inclusion of decoherence. Vol. IV, page 143

What is the gravitational field of a single quantum particle in a double-slit experiment? As Figure 39 shows, the gravitational field almost always appears in both slits, and only very rarely in just one slit. Page 204

In summary, in the strand model, the combination of gravitation and quantum theory is much simpler than was expected by most researchers. For many decades it was suggested that the combination was an almost unattainable goal. In fact, in the strand model we can almost say that the two descriptions combine naturally.

TABLE 11 Correspondences between physical systems and mathematical tangles.

Physical system	Strands	Tangle type
Vacuum	many infinite unknotted strands	unlinked
Dark energy	many fluctuating infinite strands	unlinked
Elementary vector boson	one infinite strand	a curve
Quark	two infinite strands	rational tangle
Lepton	three infinite strands	braided tangle
Meson, baryon	three or more infinite strands	composed of rational tangles
Higher-order propagating fermion	two or more infinite strands	general rational tangle
Virtual particles	open or unlinked strands	trivial tangles
Composed systems	many strands	separable tangles
Graviton	two infinite twisted strands	specific rational tangle
Gravity wave	many infinite twisted strands	many graviton tangles
Horizon	many tightly woven infinite strands	web-like rational tangle
Young universe	closed strand(s)	knot (link)

Torsion, curiosities and challenges about quantum gravity

On the one hand, the strand model denies the existence of any specific effects of *torsion* on gravitation. On the other hand, the strand model of matter describes spin with the belt trick. The belt trick is thus the strand phenomenon that is closest to the idea of torsion. Therefore, exaggerating a bit in the other direction, it could also be argued that in the strand model, torsion effects are quantum field theory effects. Ref. 197

* *

The strand model describes three-dimensional space as made of tangled strands. Several similar models have been proposed in the past.

The model of space as a *nematic world crystal* stands out as the most similar. This model was proposed by Hagen Kleinert in the 1980s. He took his inspiration from the famous analogy by Ekkehart Kröner between the equations of solid-state elasticity around line defects and the equations of general relativity. Ref. 198 Ref. 199

Also in the 1980s, the mentioned posets have been proposed as the fundamental structure of space. Various models of quantum gravity from the 1990s, inspired by spin networks, spin foams and by similar systems, describe empty space as made of extended constituents. These extended constituents tangle, or bifurcate, or are connected, or sometimes all of this at the same time. Depending on the model, the constituents are lines, circles or ribbons. In some models their shapes fluctuate, in others they don't. Ref. 157 Ref. 200

Around the year 2000, another type of Planck-scale crystal model of the vacuum has been proposed by David Finkelstein. In 2008, a specific model of space, a crystal-like network of connected bifurcating lines, has been proposed by Gerard 't Hooft. Ref. 158 Ref. 201

All these models describe space as made of some kind of extended constituents in a three-dimensional background. All these models derive general relativity from these constituents by some averaging procedure. The lesson is clear: it is *not* difficult to derive general relativity from a Planck-scale model of space. It is *not* difficult to unify gravity and quantum theory. As Luca Bombelli said already in the early 1990s, the challenge for a Planck-scale model of nature is not to derive gravity or general relativity; the challenge is to derive the other interactions. So far, the strand model seems to be the only model that has provided such a derivation.

* *

The Planck force is the force value necessary to produce a change $\hbar$ in a Planck time over a Planck length. The Planck force thus appears almost exclusively at horizons. Challenge 185 e

* *

Already in the 1990s, Leonard Susskind speculated that black holes could be formed by a single wound-up string. Strands differ from strings; they differ in the number of dimensions, in their intrinsic properties, in their symmetry properties, in the fields they carry and in the ways they generate entropy. Nevertheless, the similarity with the strand model of black holes is intriguing. Ref. 202

* *

In September 2010, two years after the strand model appeared, independent research confirmed its description of physical space, as already mentioned above. In an extended article exploring the small scale structure of space from several different research perspectives in general relativity, Steven Carlip comes to the conclusion that all these perspectives suggest the common idea that 'space at a fixed time is thus threaded by rapidly fluctuating lines'. Page 163 Ref. 155

In 2011, also independently, Marcelo Botta Cantcheff modelled space as a statistic ensemble of one-dimensional 'strings'. He explained the main properties of space, including the thermodynamic properties of black holes. Ref. 203

* *

The first version of the strand model assumed that space is not defined at the cosmic horizon, and that therefore, strand impenetrability does not hold there. The same was thought to occur at black hole horizons. The newest version of the strand model does not seem to need this exception to impenetrability. Can you explain black hole entropy without it? Challenge 186 e

* *

Page 36

The strand model also allows us to answer the question whether quantum particles are black holes: no, they are not. Quantum particles are tangles, like black holes are, but particles do *not* have horizons. As a side result, the mass of all particles is lower than a Planck mass, or more precisely, lower than a Planck mass black hole.

Strands imply that gravity is weaker than the three gauge interactions. This consequence, like the low particle mass just mentioned, is due to the different origins of gravity and gauge interactions. Gravity is due to the strand tails, whereas gauge interactions are due to the tangle cores. Thus gravity is the weakest interaction in everyday

life. The observation of the weakness of gravity at everyday and other energy scales is
Ref. 204 sometimes called the *weak gravity conjecture*. It is naturally valid in the strand model.
Page 8 The conjecture is also part of the Bronshtein cube shown in Figure 1.

* *

For an observer at spatial infinity, a black hole horizon is an averaged-out tight web of strands. What does a falling observer experience? The question will still capture the imagination in many years. Such an observer will also see strands; above all, a falling observer will never hit any singularity. The details of the fall are so involved that they are not discussed here, because the fall affects both the black hole appearance and the observer.

* *

Can black hole radiation be seen as the result of trying to tear vacuum apart? Yes and no. The answer is no, because physical vacuum cannot be torn apart, due to the maximum force principle. But the answer is also yes in a certain sense, because the maximum force is the closest attempt to this idea that can be realized or imagined.

* *

The strand model makes the point that *entanglement* and the vacuum – and thus quantum gravity – have the same nature: both are due to crossing strands. This idea
Ref. 205 has been explored independently by Mark van Raamsdonk.

* *

As we have seen, the strand model predicts no *observable* violation of Lorentz-invariance – even though it predicts its violation at Planck scale. Strands predict the lack of dispersion, birefringence and opacity of the vacuum. Strands predict that the vacuum has three dimensions whenever it is observed and that it is unique, without phase transitions. We already mentioned the impossibility of detecting single gravitons.

All these negative predictions are examples of the *'no avail'* conjecture:

▷ Quantum gravity effects cannot be distinguished from ordinary quantum fluctuations.

Despite many attempts to disprove it, all experiments so far confirm the conjecture. Because both quantum gravity effects and quantum effects are due to tail fluctuations, the strand model seems to imply the conjecture.

* *

Ref. 206 The strand model of black holes also confirms a result by Zurek and Thorne from the 1980s: the entropy of a black hole is the logarithm of the number of ways in which it could have been made.

* *

Argue that because of the strand model, no black hole can have a mass below the (corrected) Planck mass, about 11 μg, and thus that *microscopic black holes* do not exist. Can
Challenge 187 s you find a higher lower limit for the mass?

* *

Do atoms or the elementary fermions moving inside matter emit gravitational radiation, and why? The question was already raised by Albert Einstein in 1916. The strand model answers the issue in the same way as textbook physics. Elementary particles in atoms – in the ground state – do not emit gravitational waves for the same reason that they do not emit electromagnetic waves: for atoms in the ground state, there is no lower state into which they could decay. Excited atomic states do not emit gravitational waves because of the extremely low emission probability; it is due to the extremely low mass quadrupole values.

* *

Ref. 207 In 2009 Mikhail Shaposhnikov and Christof Wetterich argued that if gravitation is 'asymptotically safe', there is no physics beyond the standard model and the Higgs mass must be around 126 GeV – exactly the value that was found experimentally a few years afterwards. A quantum field theory is called *asymptotically safe* if it has a fixed point at extremely high energies. Does the strand model imply that gravity is – maybe only effectively – asymptotically safe? Challenge 188 ny

* *

Page 59 It is often stated that general relativity does not allow the description of fermions if the topology of space is kept fixed. This is wrong: the strand model shows that fermions can be included in the case that space is seen as an average of extended fundamental entities.

* *

Following the fundamental principle of the strand model, G is the fundamental constant that describes gravitation. The strand model predicts that gravity is the same for all energy scales; in other words, the constant G is *not* expected to change with energy. This agrees with recent results from quantum gravity and distinguishes the behaviour of G from that of the coupling constants in the gauge interactions of particle physics. Ref. 208

Predictions of the strand model about gravity

As just presented, the strand model makes several verifiable predictions about general relativity and quantum gravity.

— The maximum energy speed in nature is c, at all energy scales, in all directions, at all times, at all positions, for every physical observer. This agrees with observations.
— No deviations from special relativity appear for any measurable energy scale, as long as gravity plays no role. No 'double' or 'deformed special relativity' holds in nature, even though a maximum energy-momentum for elementary particles does exist in nature. Whenever special relativity is not valid, general relativity, or quantum field theory, or both together need to be used. This agrees with observations.
— There is a maximum power or luminosity $c^5/4G$, a maximum force or momentum flow $c^4/4G$, and a maximum mass change rate $c^3/4G$ in nature. The limits hold for all energy scales, in all directions, at all times, at all positions, for every physical observer. These predictions agree with observations, though only few experimental ob-

servations are close to these limit values.

— There is a minimum distance and a minimum time interval in nature. There is a maximum curvature and a maximum mass density in nature. There are no singularities in nature. All this agrees with observations, including the newly discovered black hole mergers.
— The usual black hole entropy expression given by Bekenstein and Hawking holds. The value has never been measured, but is consistently found in all calculations performed so far. In fact, black hole entropy is related to the Fulling–Davies–Unruh effect, which itself is related to the Sokolov–Ternov effect. This latter effect has already been observed in several accelerators, for the first time in 1971. However, it now seems that this observation does not actually prove black hole entropy. Vol. V, page 147
— There are no deviations from general relativity, as described by the Hilbert action, for any measurable scale. The only deviations appear in situations with a few strands, i.e., in situations where quantum theory is necessary. This agrees with observations, including those of black hole mergers, but experimental data are far from sufficient; undetected deviations could still exist.
— There is no modified Newtonian dynamics, or MOND, with evolution equations that differ from general relativity. The rotation curves of stars in galaxies are due to dark matter, to other conventional explanations, or both. Ref. 209
— There is no effect of torsion that modifies general relativity. This agrees with observations.
— There is no effect of higher derivatives of the metric on the motion of bodies. This agrees with observations, but experimental data are far from sufficient.
— Observations are independent of the precise strand fluctuations. Mathematical consistency checks of this prediction are possible.
— No wormholes, no negative energy regions and no time-like loops exist. This agrees with observations, but experimental data are far from covering every possible loophole.
— The Penrose conjecture and the hoop conjecture hold. Here, a mathematical consistency check is possible.
— There are no cosmic strings and no domain walls. This agrees with observations, but experimental data are far from exhaustive.
— Gravitons have spin 2; they return to their original state after a rotation by π and are bosons. This agrees with expectations.
— Gravitons cannot be detected, due to the indistinguishability with ordinary quantum fluctuations of the detector. This agrees with data so far.
— Atoms emit neither gravitational waves nor gravitons.
— Gravitational waves exist and can be detected. This agrees with various experiments; the final, direct confirmation occurred in late 2015.
— The gravitational constant G does not run with energy – as long as the strand diameter can be neglected. In this domain, G is not renormalized. This prediction agrees with expectations and with data, though the available data is sparse.

All listed predictions are unspectacular; they are made also by other approaches that contain general relativity as limiting cases. In particular, the strand model, like many other approaches, predicts:

Page 309

▷ With the exception of the cosmological constant and of particle masses (and possibly the Sokolov–Ternov effect), *no quantum gravity effects will be observed.*

Gravity will not yield new measurable quantum effects. So far, this prediction agrees with experiment – and with almost all proposed models of quantum foam in the research literature. In other words, we have found *no unexpected* experimental predictions from the strand model in the domain of quantum gravity. This is the so-called *'no avail'* conjecture; and it is not a surprise. Ref. 97

Page 8 In fact, the Bronshtein cube of Figure 1 also implies:

▷ There is *no* separate theory of quantum gravity that includes relativity but does not include the other interactions.

There is no room for a theory of relativistic quantum gravity in nature.

In short, strands lead us to expect deviations from general relativity only in two domains: in cosmology (such as changes of the cosmological constant) and in particle physics. The rest of this chapter deals with cosmology. The subsequent chapters focus on particle physics.

Cosmology

Cosmology is an active field of research, and new data are collected all the time. We start with a short summary.

The sky is dark at night. This and other observations about the red shift show that the universe is surrounded by a horizon and is of finite size and age. Modern measurements show that cosmic age is about 13 800 million years. The universe expands; the expansion is described by the field equations of general relativity. The universe's expansion accelerates; the acceleration appears to be described by the *cosmological constant* Λ – also called *dark energy* – that has a small positive value. The universe is observed to be flat, and, averaged over large scales, homogeneous and isotropic. At present, the observed average matter density in the universe is about 18 times smaller than the energy density due to the cosmological constant. In addition, there appears to be a large amount of matter around galaxies that does not radiate; the nature of this *dark matter* is unclear. Galaxy formation started from early density fluctuations; the typical size and amplitude of the fluctuations are known. The topology of space is observed to be simple.

The strand model, like any unified description of nature, must reproduce and explain these observations. Otherwise, the strand model is wrong.

The finiteness of the universe

In the strand model, cosmology is based on the following idea:

▷ The ***universe*** is made of *one* fluctuating strand that criss-crosses from and to the horizon. Fluctuations increase the complexity of the strand tangledness

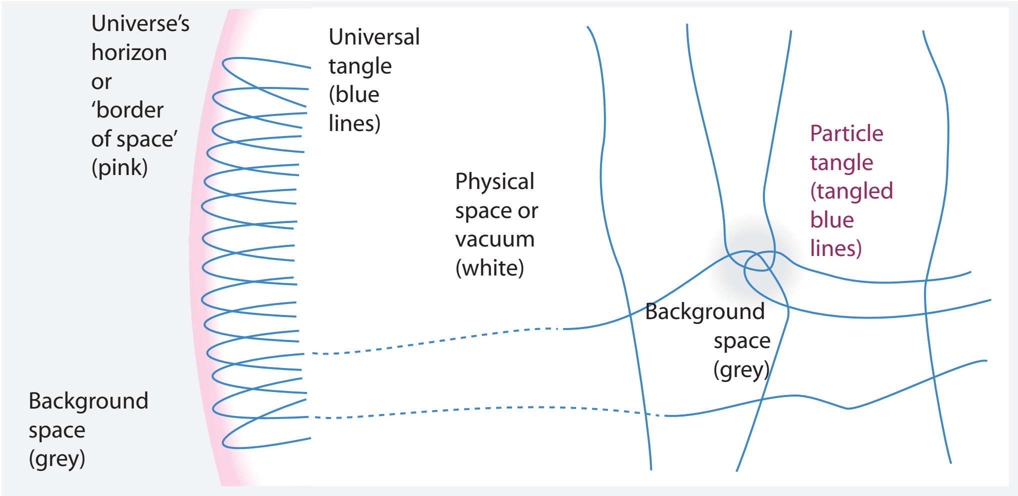


FIGURE 81 In the strand model, the universe is limited by a horizon, as schematically illustrated here. Physical space (white) matches background space (grey) only inside the horizon. Physical space thus only exists inside the cosmic horizon.

over time.

In other words, the strands of all particles are woven into the sky. The existence of finite size and of finite age then follows automatically:

▷ The ***universe's horizon*** appears at the age or distance at which the strand crossings cannot be embedded any more into a common three-dimensional background space. The horizon expands over time.

The strand model thus has a simple explanation for the finiteness of the universe and the horizon that bounds it: The universe's horizon is a weave that joins all strand tails. A schematic illustration of the cosmic horizon is given in Figure 81.

The strand model predicts that the horizon of the universe is an *cosmological particle horizon*, an event horizon similar to that of a black hole. Until 1998, this possibility Ref. 210 seemed questionable; but in 1998, it was discovered that the expansion of the universe is accelerating. This discovery implies that the cosmic horizon is indeed an event horizon, Ref. 211 as required by the strand model.

The strand model predicts that the universe is a kind of *inverted back hole*. Like for any situation that involves a horizon, the strand model thus does not allow us to make statements about properties 'before' the big bang or 'outside' the horizon. As explained above, strands predict that there is nothing behind a horizon. Page 288

In particular, the strand model implies that the matter that appears at the cosmic horizon during the evolution of the universe appears through Bekenstein–Hawking radiation. This contrasts with the 'classical' explanation form general relativity that new matter appears simply because it existed behind the horizon beforehand and then crosses the horizon into the 'visible part' of the universe.

We note that modelling the universe as a single strand implies that it contains tangles.

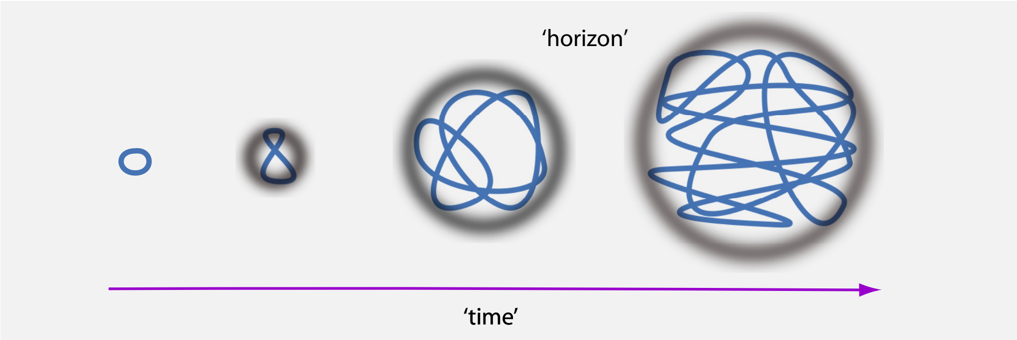


FIGURE 82 An extremely simplified view of how the universe evolved near the big bang. In this evolution, physical time, space and the surrounding horizon are in the process of getting defined.

In other words, the strand model makes the prediction that the universe cannot be empty, but that it must contain particles. Strand cosmology also confirms that the question of Page 102 initial conditions for the universe does not really make sense: particles appear at the horizon.

We also note that describing the universe as made of a single strand is a natural, but somewhat unusual way to incorporate what particle physicists and cosmologists like to call *holography*. Holography is the idea that all observables of a physical system are defined on a boundary enclosing the system. In other words, if we would know, at Planck scale, everything that happens on the walls of a room, we could know everything that is and goes on inside the room. Instead of holography, we could also call it the *NSA dream*. Holography is a consequence of the extension of the fundamental constituents of nature and is a natural consequence of the strand model. As a consequence, strand cosmology Challenge 189 e naturally reproduces holographic cosmology – though not fully, as is easy to check.

> Or cette liaison ou cet accommodement de toutes les choses créées à chacune, et de chacune à toutes les autres, fait que chaque substance simple a des rapports qui expriment toutes les autres, et qu'elle est par conséquent un miroir vivant perpétuel de l'univers.*
> Gottfried Wilhelm Leibniz, *Monadologie*, 56.

The big bang – without inflation

Any expanding, homogeneous and isotropic matter distribution had earlier stages of Ref. 212 smaller size and higher density. Also the universe has been hotter and denser in the past. Page 293 But the strand model also states that singularities do not appear in nature, because there is a highest possible energy density. As a result, the big bang might be imagined as illustrated in Figure 82. Obviously, physical space and time are not well defined near that situation, so that the figure has to be taken with a grain of salt. Nevertheless, it shows how

* 'Now this connexion or adaptation of all created things to each and of each to all, means that each simple substance has relations which express all the others, and, consequently, that it is a perpetual living mirror of the universe.'

the evolution of the universe can be seen as resulting from the increase in tangledness of the strand that makes up nature.

The strand model leads to the conjecture that the evolution of the universal strand just after the big bang automatically yields both a homogeneous and isotropic matter distribution and a flat space. Also the scale invariance of early density fluctuations seems natural in the strand model. In short, the strand model looks like a promising alternative to *inflation*: the hypothesis of inflation becomes unnecessary in the strand model, because strand cosmology directly makes the predictions that seem so puzzling in classical cosmology. This issue is still subject of research.

The cosmological constant

At present (2019), the issue of the cosmological constant, thus of dark energy, is under investigation. Watch this space for updates.

The value of the matter density

The strand model predicts that horizons emit particles. As a consequence, the strand model predicts an upper limit for the number N_b of baryons that could have been emitted by the cosmic horizon during its expansion. For a horizon shining throughout the age of the universe t_0 while emitting the maximum power $c^5/4G$, we get

$$N_{b0} \leqslant \frac{t_0\, c^5/4G}{m_b c^2} = 2.6 \cdot 10^{79} \,. \tag{194}$$

Equality would hold only if the contributions of photons, electrons, neutrinos and dark matter could be neglected. In short, using the age $t_0 = 13.8$ Ga, the strand model predicts that at most $2.6 \cdot 10^{79}$ baryons exist in the universe at present. Modern measurements indeed give values around this limit. Ref. 215

More about matter and energy densities will be added here soon.

Open challenge: What are the effects of dark matter?

Conventionally, it is argued that *cold* dark matter exists for three reasons: First, it is necessary to grow the density fluctuations of the cosmic microwave background rapidly enough to achieve the present-day high values. Secondly, it is needed to yield the observed amplitudes for the acoustic peaks in the cosmic background oscillations. Third, it explains the rotation curves observed in hundreds of galaxies and galaxy clusters. Can the strand model change these arguments? Challenge 190 ny

Later on, it will be argued that following the strand model, dark matter can only be a mixture of conventional matter and black holes. How does this dark matter prediction explain the galaxy rotation curves? This leads to a really speculative question: Could tangle effects at the scale of a full galaxy be related to dark matter? Research is ongoing, and will be added here in the coming months. Challenge 191 ny

The topology of the universe

In the strand model, physical space-time, whenever it is defined, *cannot* be multiply connected. Also all quantum gravity approaches make this prediction, and the strand model confirms it: because physical space-time is a result of averaging strand crossing switches, non-trivial topologies (except black holes) do not occur as solutions. For example, the strand model predicts that wormholes do not exist. In regions where space-time is undefined – at and beyond horizons – it does not make sense to speak of space-time topology. In these regions, the fluctuations of the universal strand determine observations. In short, the strand model predicts that all searches for non-trivial *macroscopic* (and microscopic) topologies of the universe, at both high and low energies, will yield negative results. So far, this prediction agrees with all observations.

Predictions of the strand model about cosmology

In the domain of cosmology, the strand model makes the following testable predictions.

- The universe is not empty. (Agrees with observation.)
- Its integrated luminosity does not exceed the power limit $c^5/4G$. (Agrees with observation.)
- The universe's energy density does not exceed the entropy bound. (Agrees with observation.)
- The are no singularities in nature. (Agrees with observation.)
- The matter density of the universe decreases with age, roughly as $\rho t \sim 1/t^2$. (Checks are under way. This prediction differs from the usual cosmological models.)
- There is nothing behind the cosmic horizon. Matter, energy and space appear at the horizon. (Agrees with observations and requirements of logic.)
- Early density fluctuations are scale-invariant. (Agrees with observation.)
- The universe is flat and homogeneous. (Agrees with observation.)
- Apart (maybe) from the cosmological constant Λ, all other fundamental constants of nature are constant over time and space. (Agrees with observation, despite occasional claims of the contrary.)
- Inflation is unnecessary.
- The universe's topology is trivial. There are no wormholes, no time-like loops, no cosmic strings, no toroidal black holes, no domain walls and no regions of negative energy. (Agrees with observation.)
- The above statements are independent of the precise fluctuation details. (Can be tested with mathematical investigations.)

All these predictions can and will be tested in the coming years, either by observation or by computer calculations.

SUMMARY ON MILLENNIUM ISSUES ABOUT RELATIVITY AND COSMOLOGY

We have deduced special relativity, general relativity and cosmology from the strand model. The fundamental principle of the strand model implies the invariant Planck units, the Lagrangian and action of general relativity, the finiteness of the universe and, above all, it explains in simple terms the entropy of black holes.

Space-time foam is replaced by the strand model of the vacuum: empty space is the time-average of untangled strands. More precisely, space is the thermodynamic average of crossing switches that are due to shape fluctuations of untangled strands.

The strand model – and in particular, the strand model of the vacuum – explains the number of space-time dimensions, the vacuum energy density, the matter density and the finiteness of the universe. The cosmological constant is a consequence of the finite size of the universe. The issue of the initial conditions of the universe has been defused. The macroscopic and microscopic topology of the universe is simple. And dark matter is predicted to be, as shown in the next chapter, a combination of conventional matter and black holes. Page 355

The most important predictions of the strand model are the decrease of the cosmological constant with time and the absence of inflation. Various experiments will test these predictions with increased precision in the coming years. So far, measurements do not contradict these predictions.

The strand model confirms that the speed of light c and the corrected Planck force $c^4/4G$ are *limit* values. The strand model also predicts that no variation in space and time of c, G, $\hbar$ and k can be detected, because they define all measurement units.

The strand model predicts that the cosmological constant and the masses of the elementary particles are the *only* quantum effects that will be observed in the study of gravitation. Strands strongly suggest that additional effects of quantum gravity cannot be measured. In particular, no effects of space-time foam will be observed. Page 279

The strand model is, at present, the simplest – but not the only – known model of quantum gravity that allows deducing all these results. In particular, the strands' explanation of black hole entropy is by far the simplest one known.

General relativity is an approximation of the strand model. The approximation appears when the quantum of action and, in particular, the strand diameter are neglected. General relativity and cosmology thus appear by approximating $\hbar$ as 0 in the strand model – as required by the Bronshtein cube of physics that is shown in Figure 1. Strands imply that general relativity is valid for all energies below the Planck energy. In other words, the strand model is not a generalization of general relativity. This conforms to the list of requirements for the final theory. Page 8 Page 149

If we look at the millennium list of open issues in physics, we see that – except for the issue of dark matter – all issues about general relativity and cosmology have been settled. The strand model explains the mathematical description of curved space-time and of general relativity. The strand model also provides a simple model of quantum gravity – maybe the simplest known one. Above, we had already shown that the strand model explains all mathematical structures that appear in quantum theory and in particle physics. Together with the results from this chapter we can now say: *the strand model explains* all Page 164 Page 276

concepts, i.e., all mathematical structures that appear in physical theories. In particular, strands explain the metric, curvature, wave functions, field intensities – and the probabilistic behaviour of all of them. They all result from averaging crossing switches.

In summary, starting from the fundamental principle of the strand model, we have understood that strands are the origin of gravitation, general relativity, quantum gravity and cosmology. We have also understood the mathematical description of gravitation – and, before, that of quantum physics – found in all textbooks. These results encourage us to continue our quest. Indeed, we are not done yet: we still need to deduce the possible elementary particles and to explain their properties.

CHAPTER 11

THE PARTICLE SPECTRUM DEDUCED FROM STRANDS

> No problem can withstand the assault of sustained thinking.
>
> Voltaire**

Ref. 218

Strands describe quantum theory, gauge interactions and general relativity. But do strands also settle all issues left open by twentieth-century physics? Do they settle the origin of all the elementary particles, their quantum numbers, their masses and their mixing angles? How does the infinite number of possible tangles lead to a finite number of elementary particles? And finally, do strands explain the coupling constants?
Page 164
In the millennium list of open issues in fundamental physics, these are the issues that remain. The strand model is correct only if these issues are resolved.

In this chapter, we show that the strand model indeed explains the known spectrum of elementary particles, including the three generations of quarks and leptons. The strand model is the first approach of modern physics that can provide such an explanation.

It should be stressed that from this point onwards, the ideas are particularly speculative. In the chapters so far, the agreement of the strand model with quantum field theory and general relativity has been remarkable. The following chapters assign specific tangles to specific particles. Such assignments are, by nature, not completely certain. The speculative nature of the ideas now becomes particularly apparent.

PARTICLES AND QUANTUM NUMBERS FROM TANGLES

In nature, we observe three entities: vacuum, horizons and particles. Of these, (quantum) *particles* are *localized* entities with specific *intrinsic* properties, i.e., properties that do not depend on their motion.

In nature, all the intrinsic properties of every particle, every object and every image are completely described by three types of *basic* properties: (1) the elementary particles they contain, (2) their behaviour under space-time transformations, (3) their interactions. The full list of these basic intrinsic properties of particles is given in Table 12.

Given the basic intrinsic properties for each elementary particle, physicists can deduce *all* those intrinsic particle properties that are *not* listed; examples are the half life, decay modes, branching ratios, electric dipole moment, T-parity, gyromagnetic ratio or electric polarizability. Of course, the basic intrinsic properties also allow physicists to deduce *every* property of every object and image, such as size, shape, colour, brightness, density,

** Voltaire (b. 1694 Paris, d. 1778 Paris) was an influential philosopher, politician and often satirical writer.

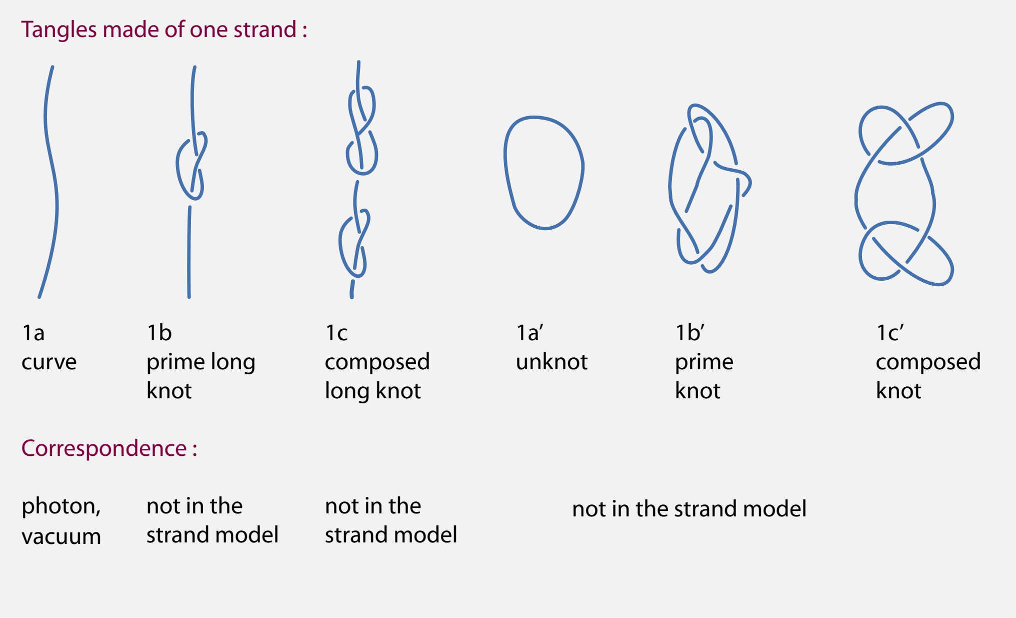


FIGURE 83 Examples for each class of tangles made of one strand.

elasticity, brittleness, magnetism or conductance.

In short, understanding *all* properties of matter and images thus only requires understanding the *basic* properties of quantum particles; and understanding the *basic* properties of quantum particles only requires understanding the *basic* properties of the *elementary* particles.

The strand model states that all elementary (and all composed) particles are tangles of strands. This leads us to ask: Which tangle is associated to each elementary particle? What kinds of elementary particles are possible? Do these tangles reproduce, for each elementary particle, the observed values of the basic properties listed in Table 12?

Page 175

It turns out that the strand model only allows a *limited number* of *elementary* particles. In addition, the tangles of these elementary particle have intrinsic properties that *match* the observed properties. To prove these strong statements, we first recall that all massive elementary particles are represented by an *infinite sequence* of tangles. We now explore tangles according to the number of strands they are made of.

PARTICLES MADE OF ONE STRAND

In the strand model, all particles made of *one* strand have spin 1, are elementary, and are bosons. Conversely, all massless elementary spin-1 bosons can only have two tails, and thus must be made of a single strand. Such one-stranded tangles return to the original strand after a core rotation by 2π. Massive elementary spin-1 bosons can have one or more strands. Tangles of more than one strand can only have spin 1 if they represent massive elementary or composed particles. In short, classifying one-stranded tangles allows classifying all elementary gauge bosons.

Page 176

TABLE 12 The full list of *basic* intrinsic properties of quantum particles, from which all other observed intrinsic properties of particles, objects and images can be deduced.

Property	Possible Value	Determines
Quantum numbers due to space-time symmetries:		
Spin S or J	integer or half-integer multiple of $\hbar$	statistics, rotation behaviour, conservation
P parity	even (+1) or odd (−1)	behaviour under reflection, conservation
C parity	even (+1) or odd (−1)	behaviour under charge conjugation, conservation
Interaction properties:		
Mass M	between 0 and the Planck mass	gravitation, inertia
Electric charge Q	integer multiples of one third of electron or proton charge	Lorentz force, coupling to photons, conservation
Weak charge	rational multiple of weak coupling constant	weak scattering and decays, coupling to W and Z, partial conservation
Mixing angles	between 0 and $\pi/2$	mixing of quarks and neutrinos, flavour change
CP-violating phases	between 0 and $\pi/2$	degree of CP violation in quarks and neutrinos
Strong charge, i.e., colour	rational multiple of strong coupling constant	confinement, coupling to gluons, conservation
Flavour quantum numbers, describing elementary particle content:		
Lepton number(s) L'	integer(s)	conservation in strong and e.m. interactions
Baryon number B	integer times 1/3	conservation in all three gauge interactions
Isospin I_z or I_3	+1/2 or −1/2	up and down quark content, conservation in strong and e.m. interactions
Strangeness S'	integer	strange quark content, conservation in strong and e.m. interactions
Charmness C'	integer	charm quark content, conservation in strong and e.m. interactions
Bottomness B'	integer	bottom quark content, conservation in strong and e.m. interactions
Topness T'	integer	top quark content, conservation in strong and e.m. interactions

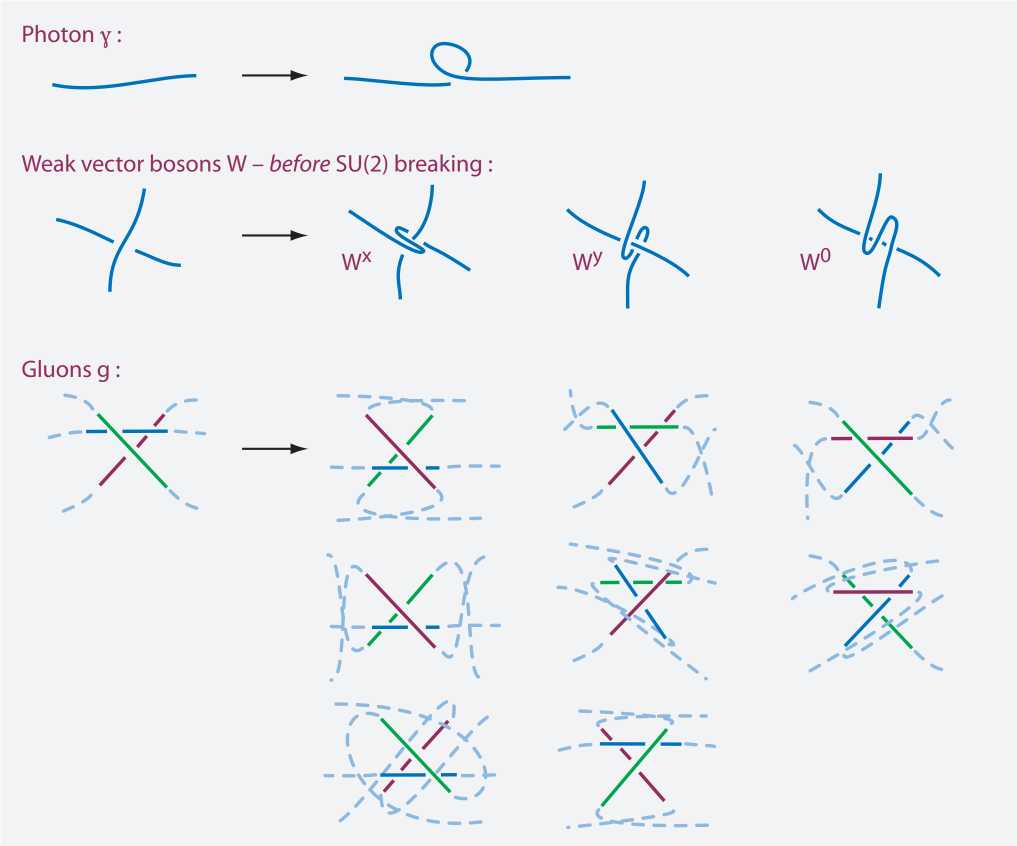


FIGURE 84 The gauge bosons in the strand model. All differ from vacuum by one curved strand – though, for clarity, the gluons are shown here using their complementary two-strand moves.

Mathematicians have already classified one-stranded tangles; they are usually called *open knots* or *long knots*. To get an overview, we list an example for each class of one-stranded tangles on the left-hand side of Figure 83. For completeness, closed curves are shown on the right-hand side of the figure. We now explore each of these classes.

Unknotted curves

The simplest type of tangle made of one strand is an *unknotted curve*, shown as example 1a in Figure 83. The study of gauge interactions has shown that unknotted strands are, Page 253 depending on their precise average shape, either vacuum strands or gauge bosons.

The time-average of a vacuum strand is straight. A single strand represents a particle if the time-averaged strand shape is not a straight line.

In the strand model, vacuum strands in flat space are, on average, *straight*. In this property, vacuum strands differ from gauge bosons, which, on average, have *curved* strands, and thus carry energy.

Gauge bosons – and Reidemeister moves

Gauge bosons are the carrier particles of the interactions. In the strand model, the gauge interactions are due to the three Reidemeister moves. The electromagnetic, the weak and the strong interaction correspond to respectively the first, second and third Reidemeister move. As we have seen above, when the three Reidemeister moves deform fermion tangle cores they generate U(1), SU(2) and SU(3) gauge symmetries. The detailed exploration Page 225 of the correspondence between tangle deformation and gauge theory led us to the gauge boson tangles shown in Figure 84.

▷ All gauge bosons – before symmetry breaking when applicable – are due to a single moving and curved strand.

A single strand represents a particle if the time-averaged strand shape is not a straight line. The lack of straightness implies non-vanishing energy. A single-strand particle can thus be either a strand with a bulge or a strand whose tails are not aligned along a straight line. The size of the bulge is related to the wavelength.

Page 228 As explained above, the *first* Reidemeister move, the twist, leads to the modelling of photons as helical strands. Therefore, photons have vanishing mass and two possible polarizations. Photons do not have tangled, localized family members; photons are massless. Their specific unknotted and twisted strand shapes also imply that photons generate an Abelian gauge theory and that photons do not interact among themselves. Automatically, photons have no weak and no strong charge. The strand model further implies that photons have negative P-parity and C-parity, as is observed.

Page 246 The study of the *second* Reidemeister move, the *poke*, showed that deformations induced by pokes can also involve braiding of tangle tails; this leads to the symmetry breaking of the weak interaction. As a result, the observed W and the Z boson strands become massive. The tangle of the W is chiral, and thus it is electrically charged; the tangle of the Z is achiral and thus electrically neutral. Being tangled, the W and the Z also carry weak charge and thus interact among themselves, generating a *non-Abelian* gauge theory. The strand model also implies that the W and the Z have no P-parity, no C-parity and no colour charge, as is observed.

Page 258 The study of the *third* Reidemeister move, the slide, led us to the existence of eight gluons. The eight gluons are unknotted, thus they carry no mass, no electric charge and no weak charge. Each gluon tangle has two possible polarizations. The strand model of gluons also implies that they have negative P-parity and no C-parity, as is observed. Gluons tangles carry colour and interact among themselves, thus they generate a non-Abelian gauge theory. In contrast to the other two interactions, free, single gluons are short-lived, because their structure induces rapid hadronization: when gluons act on the vacuum, quark–antiquark pairs are produced. Gluons do not have tangled family members; they are massless in the high energy limit, when their tails are aligned.

For completeness we mention that by assignment, all gauge bosons differ from vacuum by a single curved strand, have vanishing lepton and baryon numbers, and thus also lack all flavour quantum numbers. All this is as observed.

The strand model explains the lack of *classical SU(2) field waves* as a consequence of the breaking of the SU(2) symmetry and the consequent mass of the weak bosons.

Strands explain the lack of *classical SU(3) waves*, also called *gluonic waves*, as a consequence of the topological impossibility to produce such waves, which is related to the infinite mass of single free gluons.

In somewhat sloppy language we can say that the shape and the effects of photons are one-dimensional, those of the unbroken weak bosons are two-dimensional, and those of the gluons are three-dimensional. This is the essential reason that they reproduce the U(1), SU(2) and SU(3) groups, and that no higher gauge groups exist in nature.

In summary, Reidemeister's theorem implies that the list of known gauge bosons with spin 1 is complete. But the list of possible tangles made of a single strand is much more extensive; we are not done yet.

Open or long knots

Page 254 Single strands could also contain knotted regions. We have explained earlier on that all such possibilities – mathematically speaking, all so-called *open knots* or *long knots* – have no relation to particles. In the strand model, they cannot appear and thus play no role. The original strand model from 2008 did include such configurations as particles (for example as W and Z bosons), but it now – i.e., after 2014 – seems that this inclusion is an unnecessary complication.

Closed tangles: knots

Figure 83 shows, on the right-hand side, examples for all classes of *closed* tangles of one strand, i.e., of tangles *without tails*. They are usually just called *knots* in mathematics. In the strand model knots do not appear. They do not seem to have physical relevance and we do not explore them here.

Summary on tangles made of one strand

In summary, a single strand represents a particle if the strand shape is, on average, not a straight line. This distinguishes a vacuum strand from a particle strand. A particle strand can thus be a strand with a bulge or a strand whose tails are not aligned along a straight line. All tangles made of *one open strand* represent *elementary* particles of spin 1, thus elementary vector bosons.

Massless elementary spin-1 particles are made of one open strand also because other tangles cannot reproduce both zero mass and the spin-1 behaviour under rotations: only one-stranded tangles return to the original strand after a core rotation by 2π and allow vanishing mass at the same time.

In the strand model, the tangle made of one curved strand is assigned to the *photon*. The strand model correctly reproduces and thus explains the properties of the photon.

Particles made of two strands

In the strand model, particle tangles can also be made of *two* strands. Examples for all the classes of two-stranded tangles are given in Figure 85. Each class has a physical particle assignment.

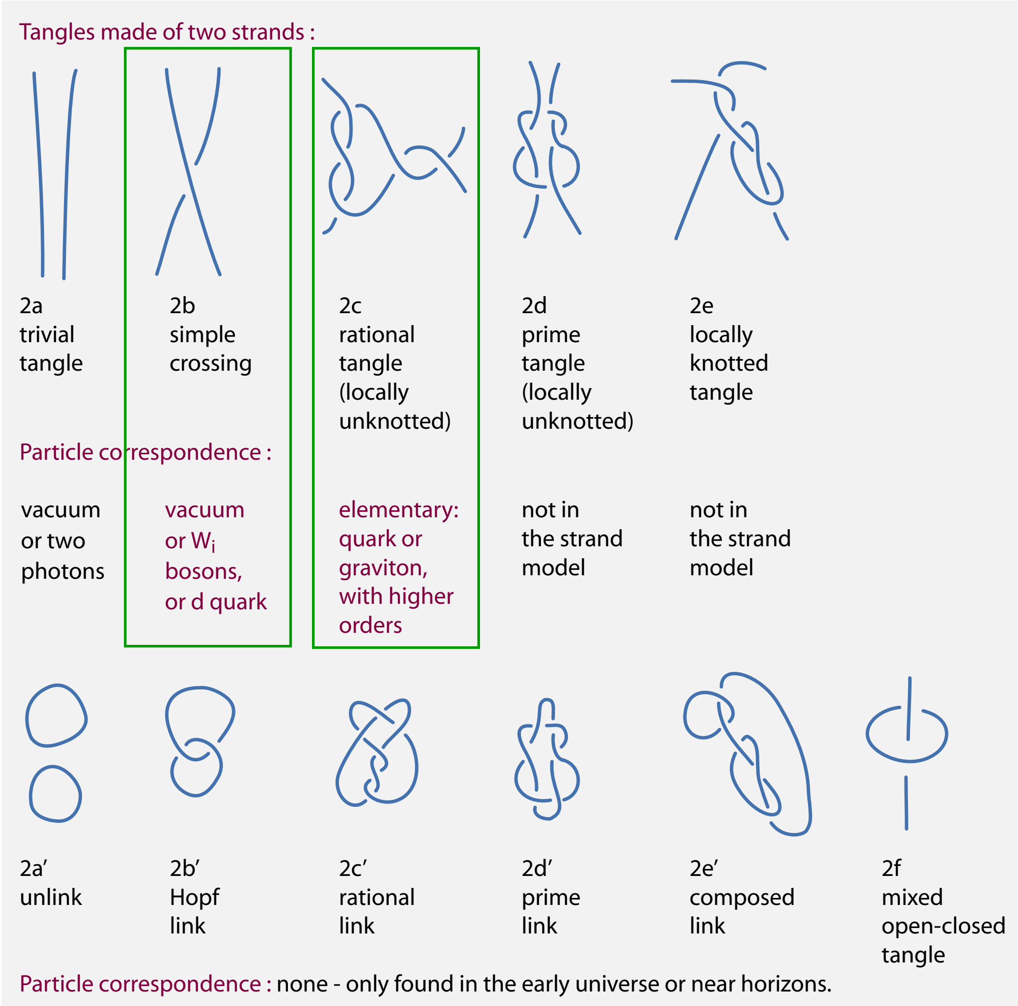


FIGURE 85 Possible tangles made of two strands.

— The simplest tangle made of two strands is the *trivial tangle*, shown as example 2a in Figure 85. In the strand model, the trivial tangle, like all *separable* tangles, is a *composite* system. Each of the two strands can represent either the vacuum or a photon. Simply stated, the trivial tangle of two strands is not an elementary particle.
— The simplest non-trivial tangle made of two strands is the *crossing*, shown as 2b in Figure 85. In the strand model, the crossing appears as part of the vacuum, or as unbroken W bosons; in addition, for certain tail configurations, it can represent a down quark, as we will see below.
— A new class of tangles are the *rational tangles*, represented by example 2c in the figure. A rational tangle is a tangle that can be untangled by moving its tails *around*. (Also example 2b is a rational tangle.) Rational tangles are distinct from prime and from locally knotted tangles, shown as examples 2d and 2e, which require pulling the tail *through* the tangle to untangle it. Rational tangles are thus *weakly* tangled. As we will

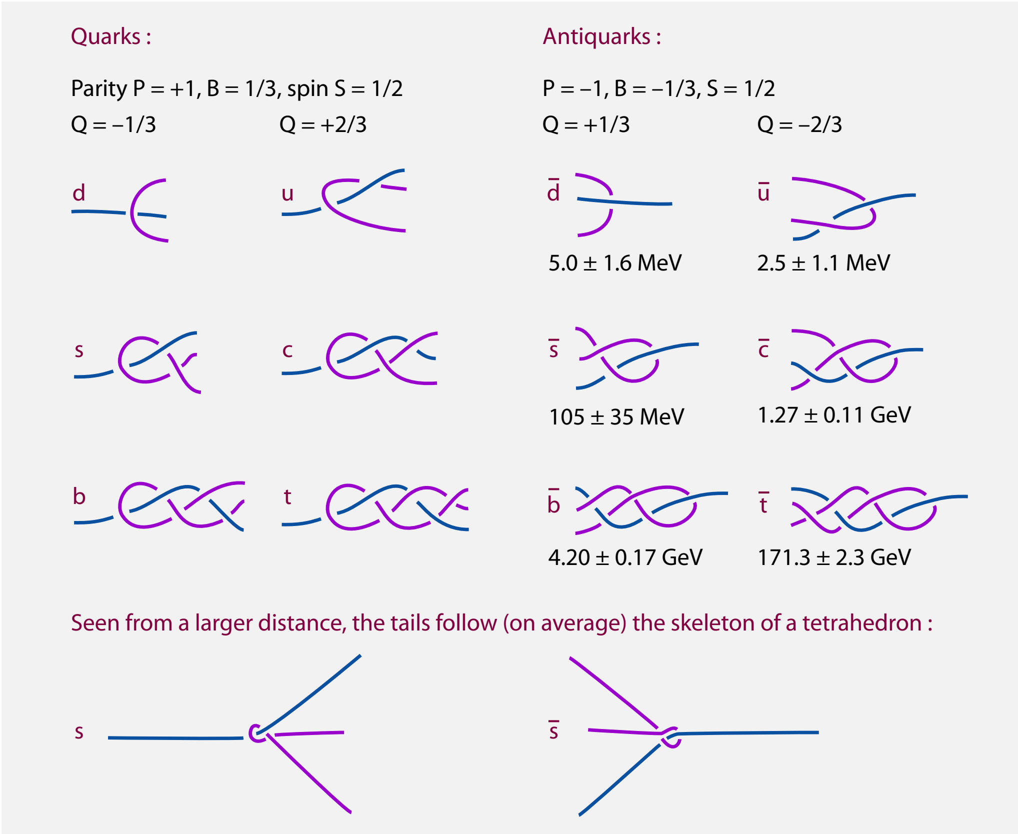


FIGURE 86 The simplest tangles assigned to the quarks and antiquarks. For reference, the experimental mass values are also given.

see,

▷ Rational tangles of two strands represent the *graviton* and the *quarks*.

We will discuss them in detail in the next two sections. More complicated rational tangles are higher-order propagating states of the simpler ones.

— Another class of tangles are *prime tangles*, for which the tangle 2d is an example. Like knotted one-stranded tangles, we conclude that prime tangles are not part of the strand model.
— Still another class of tangles are *locally knotted tangles*, shown as example 2e. Also this class is not part of the strand model.
— Finally, *closed tangles*, *links* and *mixed tangles*, shown in the lower row of Figure 85, have no role in the strand model.

In short, the only two-stranded tangles of interest in the strand model are the rational tangles. We now explore them in more detail.

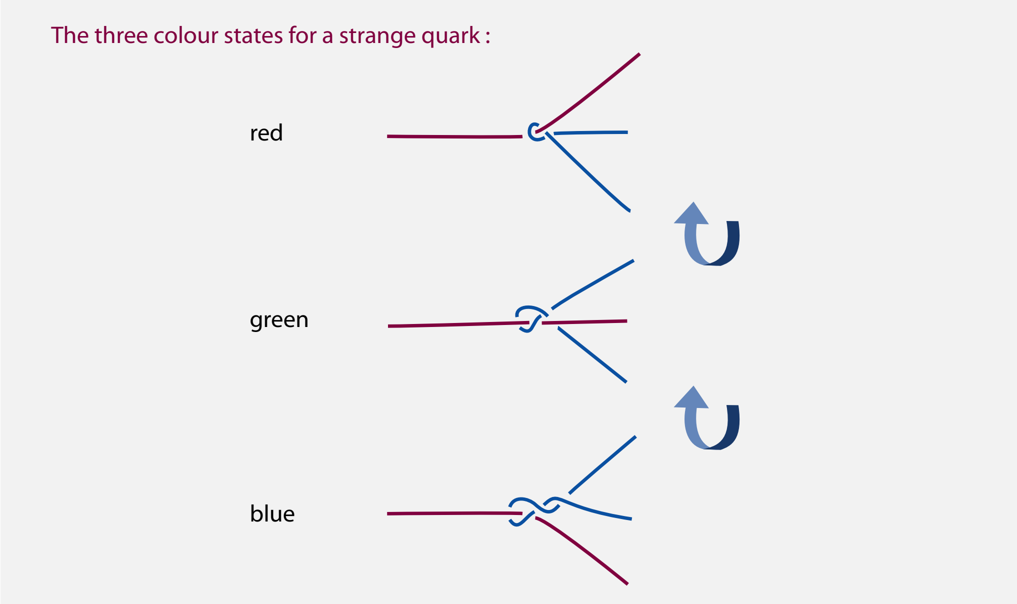


FIGURE 87 The three colour charges correspond to the three possible spatial orientations; the centre tail on the right is always above the paper plane, the other two tails on the right are below the paper plane.

Quarks

Page 271 The exploration of the strand model and of the strong interaction showed: the tangle of a coloured fermion, thus of a quark, must be rational, must reproduce the three possible colour options, and must break the three-belt symmetry.

The simplest tangles that realize these requirements are shown in Figure 86: quark tangles are *rational tangles* made of *two* strands. Higher quark generations have larger crossing numbers. The four tails form the skeleton of a tetrahedron. A particle with two strands tangled in this way automatically has spin 1/2. The electric charges of the quarks are 1/3 and −2/3, an assignment that is especially obvious for up and down quarks and that will become clearer later on, in the study of hadrons. Parity is naturally assigned as done in Figure 86. Baryon number and the other flavour quantum numbers – isospin, strangeness, charm, bottomness, topness – are naturally assigned as usual. The flavour quantum numbers simply 'count' the number of corresponding quark tangles. Like all localized tangles, quarks have weak charge. We will explore weak charge in more detail Page 382 below. Antiquarks are mirror tangles and have opposite quantum numbers. We will see Page 332, page 340 below that these assignments reproduce the observed quantum numbers of all mesons and baryons, as well as all their other properties.

We note that the simplest version of the down quark is a simple crossing; nevertheless, it differs from its antiparticle, because the simple crossing mixes with the braid with seven crossings, 13 crossings, etc.; this mixing is due to the leather trick, as shown below. Page 323 And for every quark type, these more complicated braids differ from those of their antiparticles.

For each quark, the four tails form the skeleton of a tetrahedron. In Figure 86 and

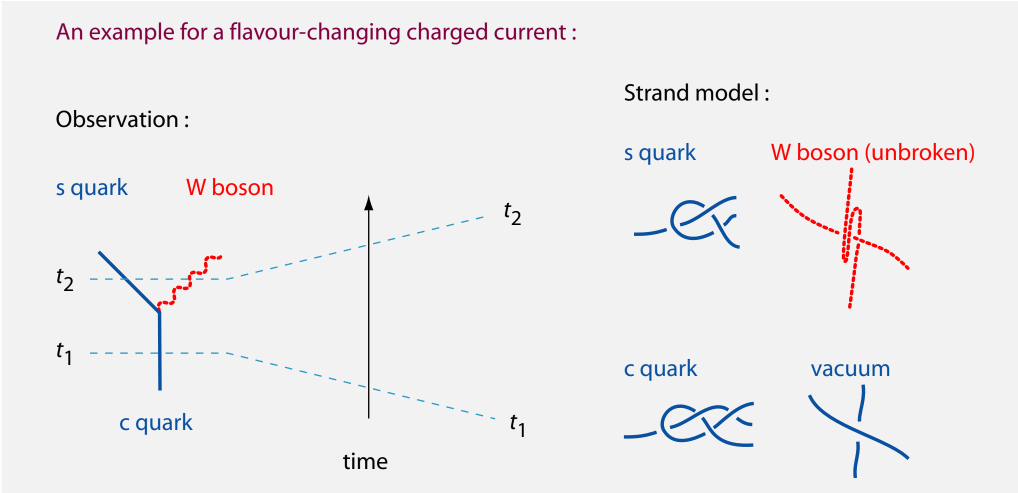


FIGURE 88 Absorption or emission of a W boson changes quark flavour.

Figure 87, the tetrahedral skeletons are drawn with one tail in the paper plane; of the other three tails, the middle one is assumed to be *above* the paper plane, and the outer two tails to be *below* the paper plane. This is important for the drawing of quark compounds later on. The three tails allow us to reproduce the strong interaction and the colour charge of the quarks: each colour is one of three possible orientations in space; more precisely, the three colours result from the three possible ways to map a quark tangle to the three Page 267 belt structure. Each colour corresponds to a different choice for the tail that lies above the paper plane, as shown in Figure 87. The colour interaction of quarks will be clarified Page 332 in the section on mesons.

In the strand model, the quark tangles thus carry *colour*. In nature, no free coloured particle has been observed. The strand model reproduces this observation in several ways. First of all, all leptons and baryons are colour-neutral, as we will see shortly. Secondly, only free quark tangles, as shown in Figure 86, have a definite colour state, because they have a fixed orientation in space. Thirdly, free quark states, thus quark states in the tetrahedral configuration of Figure 86, do not fit into vacuum even at large distances from the core; thus free quarks carry infinitely high energy. In practice, this means that free quark states do not occur in nature. Indeed, a free, coloured quark tangle can reduce its energy by interacting with one or several other quarks. The result is a strong colour attraction between quarks that leads to colourless composites.

In short, also in the strand model, only colourless composites of quarks exist as stable Page 332 free particles. We will explore quark composites and the issue of confinement of quarks Page 337 in more detail shortly.

In nature, quarks are weakly charged and interact with W bosons. In the strand model, the absorption or the emission of a W boson is the operation that takes a quark tangle and adds or subtracts a braiding step. This process is illustrated in Figure 88, which shows that a braiding (unbraiding) operation corresponds to the emission (absorption) of an W boson before symmetry breaking. It is straightforward to check that this operation fulfils all conservation laws and properties that are observed for these so-called

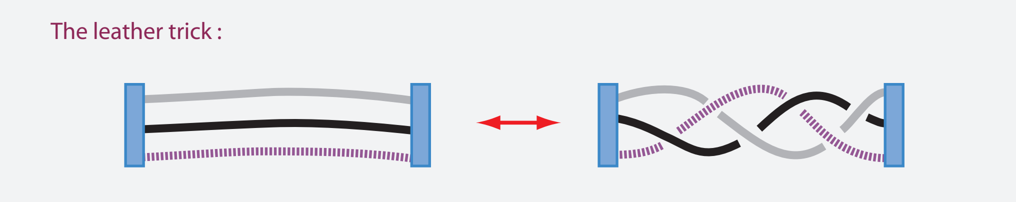


FIGURE 89 The leather trick is the deformation process that changes these two structures into each other. The leather trick limits structures made of three-stranded braids to six basic types.

flavour-changing charged currents. The absorption or emission of an (unbroken) Z boson has no braiding effect. The strand model thus reproduces the result that only the charged weak bosons can change quark flavours, as is observed.

For completeness, we mention that quarks, being tangles of *two* strands, have vanishing lepton number. Indeed, as we will see below, lepton tangles are made of *three* strands. Page 325

In summary, all quantum numbers of quarks are reproduced by the strand model, as long as quarks are modelled as braids of two strands with ends directed along the corners of a tetrahedron.

Quark generations

We stress that the quark tangles shown Figure 86 represent only the *simplest* tangle for each quark. First of all, longer braids are mapped to each of the six quarks. This might seem related to the *leather trick* shown in Figure 89. This trick is well-known to all people in the leather trade: if a braid of three strands has $n \geqslant 6$ crossings, it can be deformed into a braid with $n - 6$ crossings. We might conjecture that, due to the leather trick, there is no way to introduce more than 6 quarks in the strand model.

In fact, the leather trick argument assumes that the braid end – and thus the ends of the strands – can be moved *through* the braids. In the strand model, this can only happen at the horizon, the only region where space (and time) are not well-defined, and where such manipulations become possible. The low probability of such a process will be important in the determination of quark masses. Page 365

Instead of resting on the leather trick, it is simpler to assume that braids with large numbers of crossings are mapped modulo 6 to the braids with the smallest number of crossings. This is consistent, because in the strand model, a braid with six additional crossings is mapped to a particle together with a virtual Higgs boson. The modulo 6 rule thus represents the Yukawa mass generation mechanism in the strand model. Page 330

In summary, in the strand model, each quark is not only represented by the tangles shown in Figure 86, but also by tangles with 6 additional crossings, with 12 additional crossings, etc.

As a mathematical check, we can also ask whether *all* other rational tangles are mapped to quarks. Rational tangles of higher complexity arise by repeatedly twisting any pair of tails of a quark tangle. This process produces an infinite number of complex two-stranded tangles. In the strand model, these tangles are quarks surrounded by virtual particles. Equivalently, we can say that all the more complex rational tangles that do not appear in Figure 86 are higher-order propagators of quarks. Challenge 192 e

FIGURE 90 The graviton in the strand model.

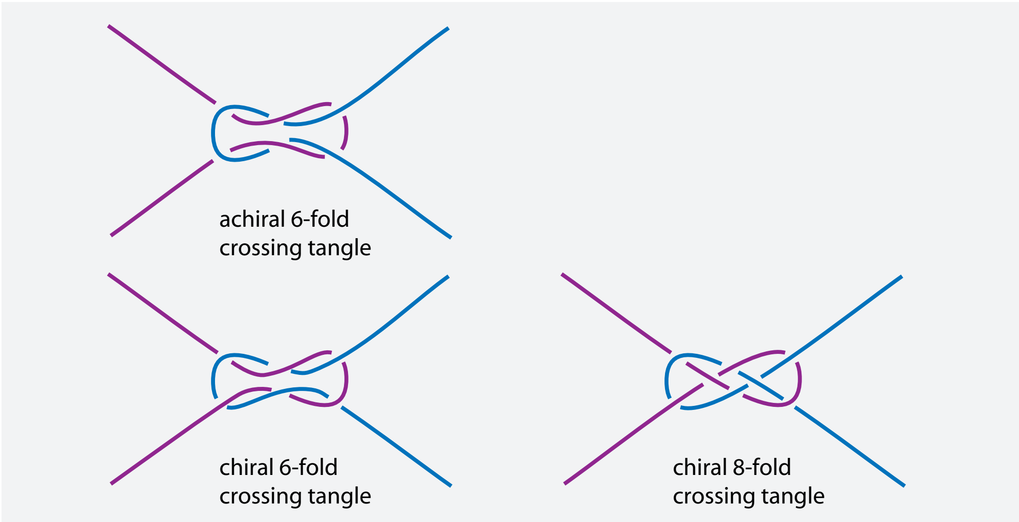


FIGURE 91 Which particle states are described by these tangles?

The graviton

Page 298 One rational tangle made of two strands is special. This special tangle is shown (again) in Figure 90. It differs from a quark tangle in one property: the tails are parallel (and near) to each other, and thus lie (almost completely) in a plane. Its tangle core returns to its original state after rotation by π, and therefore models a spin-2 particle. The tangle is not localized along its propagation direction; thus it has no mass, no electric and no weak charge. It also has no colour charge. *The tangle represents the graviton.* Similar tangles with higher winding numbers represent higher orders in the perturbation theory of gravitation.

Page 281 The chapter on gravitation has already shown how gravitons lead to curvature, horizons and the field equations of general relativity.

A puzzle

The topic of two-stranded tangles also requires to solve the puzzle of Figure 91. To which physical states do the three pictured tangles correspond? Challenge 193 s

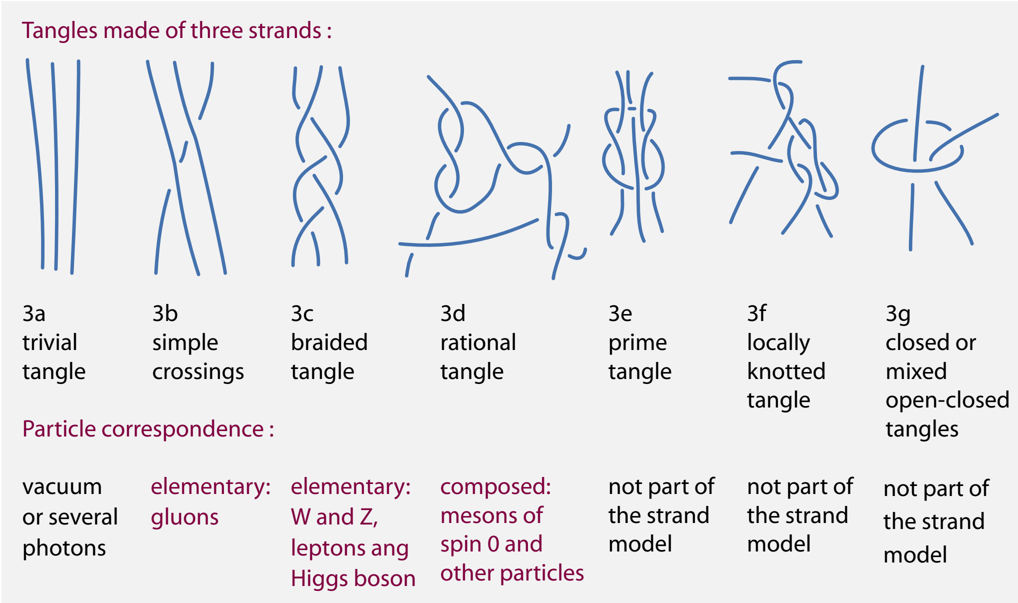


FIGURE 92 Examples for all the classes of tangles made of three strands.

Summary on two-stranded tangles

In summary, the strand model predicts that apart from the six quarks, the graviton, and the unbroken weak bosons W_1, W_2 and W_3, no other two-stranded elementary particle exists in nature.

Quarks and the graviton, the elementary particles made of two strands, are *rational* tangles. Their strand models are thus not tangled in a complicated way, but tangled in the *least complicated* way possible. This connection will be of importance in our search for elementary particles that are still undiscovered.

Particles made of three strands

In the strand model, the next group are particles made of *three* strands. Examples for all classes of three-stranded tangles are given in Figure 92. Several classes of three-stranded tangles turn out to be composites of two-stranded particles. However, a number of tangles are new and represent elementary particles.

Leptons

The candidate tangles from 2008 for the leptons shown in Figure 93 are the simplest possible non-trivial tangles with three strands. These lepton tangles are simple braids with tails reaching the border of space. The six tails probably point along the coordinate axes. These braided tangles have the following properties.

— Each lepton is localized. Each lepton has mass: its three tails con be braided, thus have non-vanishing Yukawa coupling, thus generate mass. And each lepton has spin

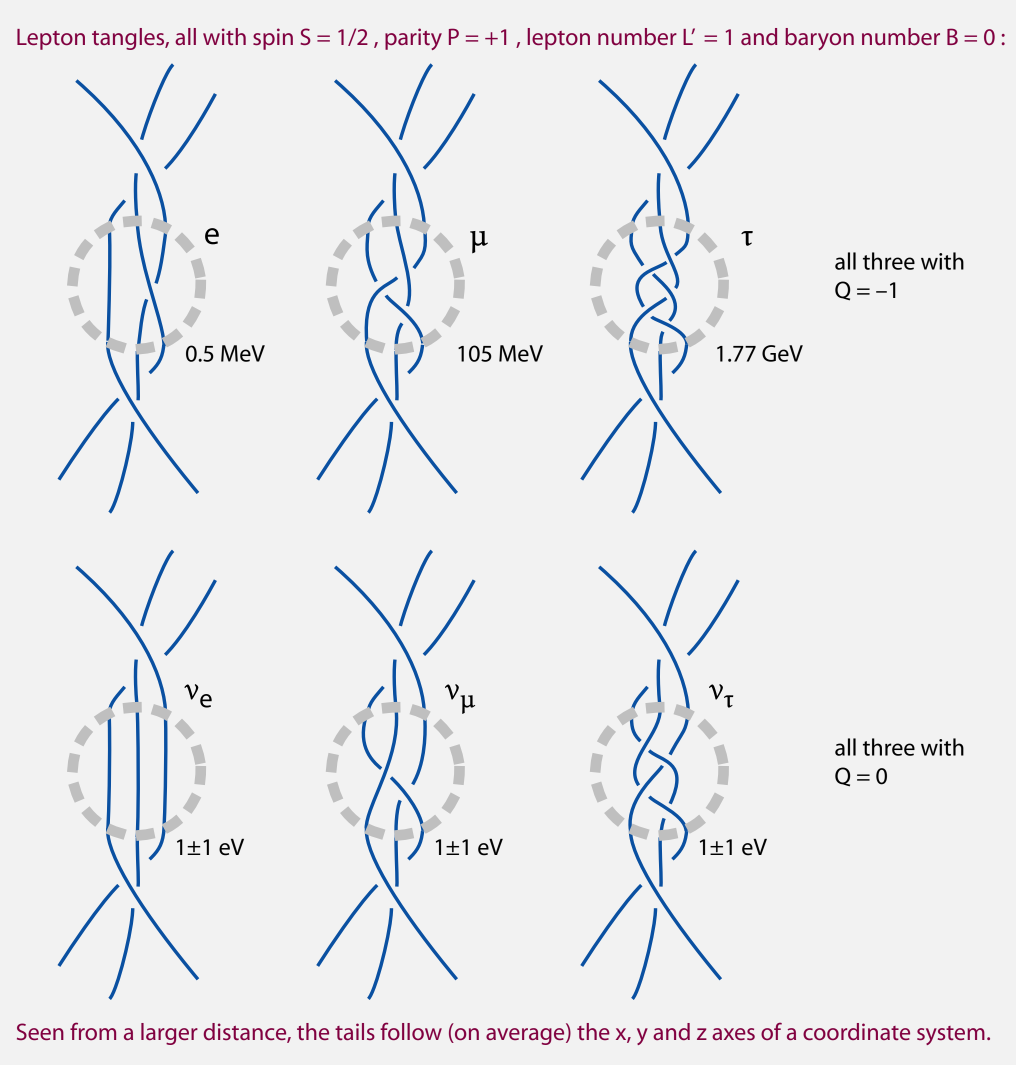


FIGURE 93 The simplest tangles of the leptons, with the experimental mass values. Antileptons are mirror tangles.

1/2. Each lepton thus follows the Dirac equation.
— Each lepton has weak charge.
— Charged leptons and antileptons differ. Each has two possible chiralities.
— Three of the tangles are topologically chiral, thus electrically charged, and three other tangles are topologically achiral, thus uncharged.
— The spatial parity P of the charged lepton tangles is opposite to that of their antiparticles.
— Being made of three strands, lepton tangles have vanishing colour charge and vanishing baryon number.
— In contrast to quarks, lepton tangles can be inserted in the vacuum using a localized, i.e., finite amount of energy and are thus predicted to exist as free particles.

— The three types of lepton (flavour) numbers can be assigned as usual; the lepton numbers are conserved in reactions, apart for neutrino mixing effects, as we will see below.
— The strand model predicts that the electron, the charged tangle with the lowest mass, is stable, as there is no way for it to decay and conserve charge and spin. The other two generations are predicted to be unstable, due to weak decays that simplify their topology.
— The three generations are reproduced by the strand model, as every more complicated braid can be seen as equivalent to one of the first six braids, with the same braiding argument that limits the number of quarks.
— There is a natural mapping between the six quarks and the six leptons. It appears when the final bend of the 'longer' quark strand is extended to the border of space, thus transforming a two-stranded quark braid into a three-stranded lepton braid. Thus we get three common generations for quarks and leptons.
— The neutrino strands differ by tail braiding; the strand model thus predicts that the weak interaction mixes neutrinos.
— All lepton tangles differ from each other. Thus the mass values are different for each lepton.
— Due to the small amount of tangling, the strand model predicts that the masses of the leptons are much smaller than those of the W and Z boson. This is indeed observed. (This also suggests a relation between the mass and the total curvature of a tight tangle.)
— The simplest tangle for the electron neutrino also suggests that the mass values for the electron neutrino is naturally small, as its tangle is almost not tangled.
— The strand model predicts that lepton masses increase with the generation number. Since the neutrino masses are not precisely known, this prediction cannot yet be checked.
— Neutrinos and antineutrinos are both massive and differ from each other. If the tangle of the electron neutrino is correct, the electron neutrino of opposite chirality is expected to be seen only rarely – as is observed.

In summary, tangles of three strands have precisely the quantum numbers and most properties of leptons. In particular, the strand model predicts exactly three generations of leptons, and predicts that all leptons have mass.

Ref. 224 This implies that searches for the neutrino-less double beta decay should yield negative results, that the magnetic moments of the neutrinos should have the exceedingly Ref. 225 small values predicted by the standard model of particle physics, and that rare muon and other decays should occur at the small rates predicted by the standard model.

Open issue: are the lepton tangles correct?

The argument that leads to the lepton tangles is vague. The tangle assignments might need corrections. There are two issues.

First, there is an aesthetic issue: in most particle tangles, the electric charge unit is Page 390 given by *three* crossings of the same sign. It seems odd that leptons should form an exception.

Secondly, the candidate tangles suggest that the muon neutrino is more massive than

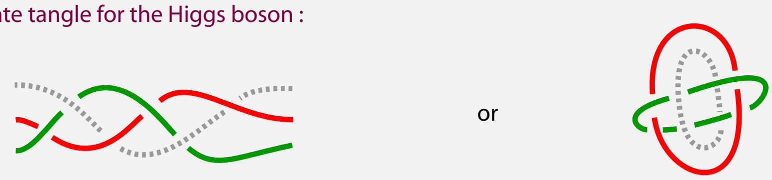


FIGURE 94 A candidate tangle for the Higgs boson in the strand model: the open version (left) and the corresponding closed version (right). For the left version, the tails approach the six coordinate axes at infinity.

the electron. Most probably therefore, the tangles need amends. Can you improve the situation, either by finding better tangles or by finding better arguments?

Challenge 194 ny

The Higgs boson – the mistaken section from 2009

The existence of the Higgs boson is predicted from the standard model of elementary particle physics using two arguments. First of all, the Higgs boson prevents unitarity violation in longitudinal W–W and Z–Z boson scattering. Secondly, the Higgs boson confirms the symmetry breaking mechanism of SU(2) and the related mass generation mechanism of fermions. Quantum field theory predicts that the Higgs boson has spin 0, has no electric or strong charge, and has positive C and P parity. In other words, the Higgs boson is predicted to have, apart from its weak charge, the same quantum numbers as the vacuum.

In the strand model, there seems to be only one possible candidate tangle for the Higgs boson, shown on the left of Figure 94. The tangle has positive C and P parity, and has vanishing electric and strong charge. The tangle also corresponds to the tangle added by the leather trick; it thus could be seen to visualize how the Higgs boson gives mass to the quarks and leptons. However, there are two issues with this candidate. First, the tangle is a deformed, higher-order version of the electron neutrino tangle. Secondly, the spin value is not 0. In fact, there is no way at all to construct a spin-0 tangle in the strand model. These issues lead us to reconsider the arguments for the existence of the Higgs boson altogether.

Page 253 We have seen that the strand model proposes a clear mechanism for mass generation:

▷ ***Mass*** is due to strand braiding.

This mechanism, due to the weak interaction, explains the W and Z boson mass ratio, Page 360 as we will see below. The leather trick that explains fermion masses can be seen as the addition of a sixfold tail braiding. In particular, the rarity of the braiding process explains why particle masses are so much smaller than the Planck mass. In short, the strand model explains mass *without* a Higgs boson.

If the Higgs boson does not exist, how is the unitarity of longitudinal W and Z boson scattering maintained? The strand model states that interactions of tangles in particle collisions are described by deformations of tangles. Tangle deformations in turn are de- Page 210 scribed by unitary operators. Therefore, the strand model predicts that unitarity is never

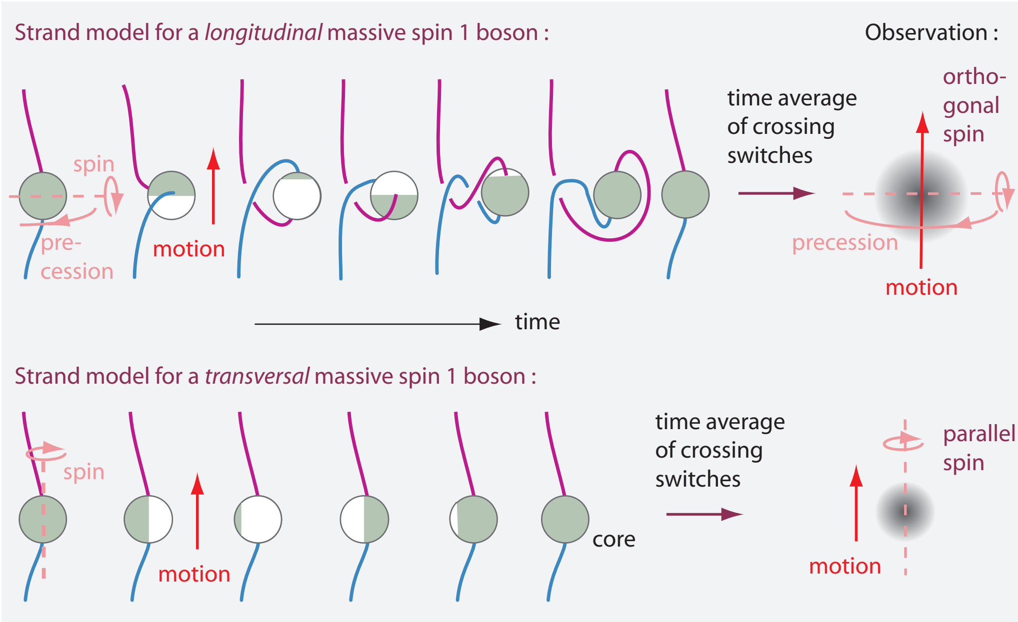


FIGURE 95 In the strand model, transverse and longitudinal W and Z bosons differ. (Note added in 2012: this statement is mistaken.)

violated in nature. In particular, the strand model automatically predicts that the scattering of longitudinal W or Z bosons does *not* violate unitarity.

In other terms, the strand model predicts that the conventional argument about unitarity violation, which requires a Higgs boson, must be wrong. How can this be? There are at least two loopholes available in the research literature, and the strand model realizes them both. Ref. 226

The first known loophole is the appearance of non-perturbative effects. It is known for a long time that non-perturbative effects can mimic the existence of a Higgs boson in usual, perturbative approximations. In this case, the standard model could remain valid at high energy without the Higgs sector. This type of electroweak symmetry breaking would lead to longitudinal W and Z scattering that does not violate unitarity. Ref. 227

The other loophole in the unitarity argument appears when we explore the details of the longitudinal scattering process. In the strand model, longitudinal and transverse W or Z bosons are modelled as shown in Figure 95. For longitudinal bosons, spin and its precession leads to a different situation than transversal bosons: longitudinal bosons are *more delocalized* than transversal bosons. This is not the case for fermions, where the belt trick leads to the *same* delocalization for longitudinal and transverse polarization. Interestingly, it is also known for a long time that different delocalization for longitudinal and transversal bosons *maintains* scattering unitarity, and that in the case of delocalization the conventional argument for the necessity of the Higgs boson is wrong. These are well-known consequences of the so-called *non-local regularization* in quantum field theory. The strand model thus provides a specific model for this non-locality, and at the same time explains why it *only* appears for longitudinal W and Z bosons. Ref. 228

The issue of different scattering behaviour for longitudinal and transverse weak bosons also raises the question whether the mass of the longitudinal and the transversal bosons are precisely equal. The possibility, triggered by Figure 95, might seem appealing at first sight in order to solve the unitarity problem. However, the strand model forbids such a mass difference. In the strand model, mass is due to tangle fluctuations, but does not depend on spin direction.

In other words, the strand model predicts that the scattering of longitudinal W and Z bosons is the first system that will show effects specific to the strand model. Such precision scattering experiments might be possible at the Large Hadron Collider in Geneva. These experiments will allow checking the *non-perturbative effects* and the *regularization effects* predicted by the strand model. For example, the strand model predicts that the wave function of a longitudinal and a transversally polarized W or Z boson of the same energy differ in cross section.

In summary, the strand model predicts well-behaved scattering amplitudes for longitudinal W and Z boson scattering in the TeV region, together with the absence of the Higgs boson.* The strand model explains mass generation and lack of unitarity violations in longitudinal W or Z boson scattering as consequences of tail braiding, i.e., as non-perturbative and non-local effects, and not as consequences of an elementary spin-0 Higgs boson. The forthcoming experiments at the Large Hadron Collider in Geneva will test this prediction.

The Higgs boson – the corrected section of 2012

In July 2012, CERN researchers from two different experiments announced the observation of a new neutral boson with a mass of 125 GeV. Additional data analysis showed Vol. V, page 254 that the boson has spin 0 and positive parity. All experimental checks confirm that the boson behaves like the Higgs boson predicted in 1963 by Peter Higgs and a number of Vol. V, page 259 other researchers.

The results lead to question several statements made in 2009 in the previous section.

— Is the tangle on the left-hand side of Figure 94 really a higher order version of the electron neutrino? It turns out that this statement is wrong: in contrast to the tangle of the neutrino, the tangle of Figure 94 is not twisted.
— Does the tangle of Figure 96 have spin 1/2 or spin 0? As mentioned already in 2009, an effective spin 0 might be possible, in a similar way that it is possible for spin-0 mesons. Spin 0 behaviour might appear because the tangle can be oriented in different directions or because of the Borromean property: no two strands have more crossings than two vacuum strands; the time average of these situations has the same symmetry as the vacuum, and thus implies spin 0.
— Does the tangle of Figure 96 have the correct, positive, C and P values expected for a Higgs boson? It seems so.
— Is the mentioned non-locality effect for W and Z bosons real? If the effect were real, it should also appear for other spin-1 particles. In the strand model, mass values should

* If the arguments against the Higgs boson turn out to be wrong, then the strand model might be saved with a dirty trick: we could argue that the tangle on the left-hand side of Figure 94 might effectively have spin 0. In this case, the ropelength of the Borromean rings, 29.03, together with the ropelengths of the weak bosons, lead to a Higgs mass prediction, to first order, in the range from $(29.03/10.1)^{1/3} \cdot 80.4\,\mathrm{GeV} = 114\,\mathrm{GeV}$ to $(29.03/13.7)^{1/3} \cdot 91.2\,\mathrm{GeV} = 117\,\mathrm{GeV}$, plus or minus a few per cent.

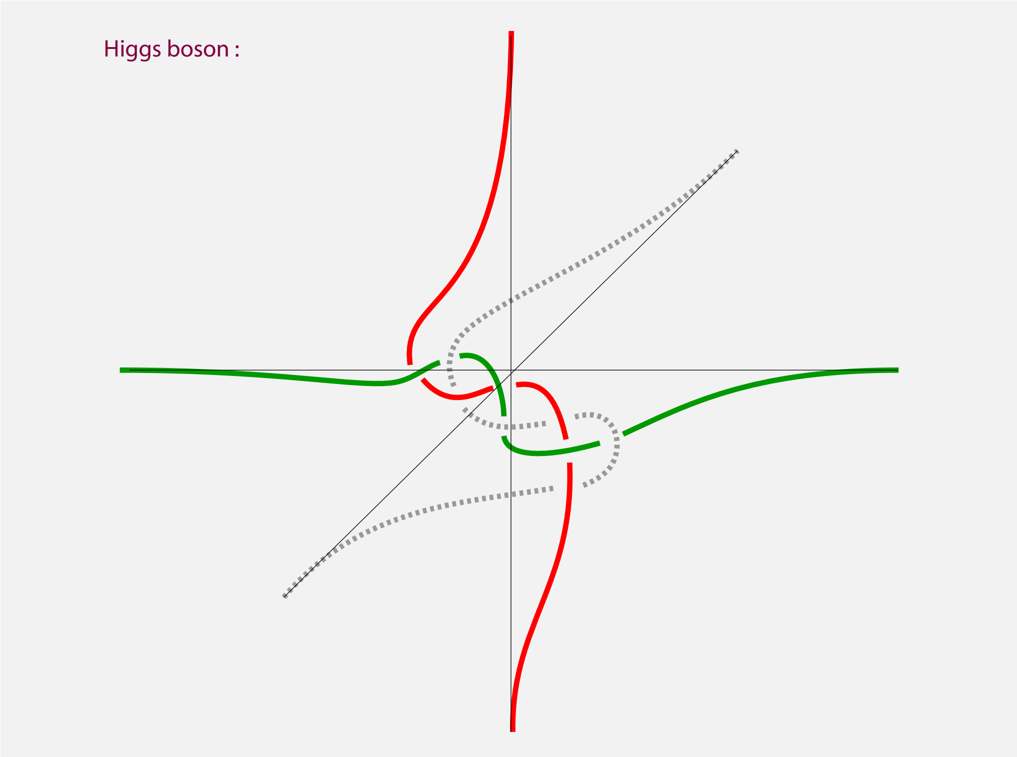


FIGURE 96 The tangle of the Higgs boson in the strand model. Spin 0 appears because the braid can be oriented in different directions, so that the time average has spherical symmetry. The tangle has 9 crossings: 3 crossings appear already in the vacuum configuration of three strands, and the additional 6 crossings (see Figure 94) are due to the Higgs boson.

not depend on spin orientation, but only on tangle core topology. The statements made in 2009 on delocalization and longitudinal scattering seem wrong in retrospect.

— Would the Higgs boson tangle assignment of Figure 96 be testable? Yes; any tangle assignment must yield the observed mass value and the observed branching ratios and decay rates. This is a subject of research. But already at the qualitative level, the proposed tangle structure of the Higgs boson suggests decays into leptons that are similar to those observed at CERN.
— Is the tangle of Figure 96 elementary? Yes.
— Are there other possible Higgs boson tangles? This issue is open. The braid structure seems the most appealing structure, as it embodies the effect of tail braiding, an effect that is important for the appearance of mass. Page 370
— Are knots and links, i.e., closed tangles, really forbidden? The discussion about the Higgs boson concerns the open tangle shown in Figure 96, not the Borromean link shown on the right-hand side of Figure 94. So far, there is no evidence for closed tangles in the strand model. Such evidence would mean a departure from the idea that nature is a single strand.
— Does the Higgs boson issue put into question the strand model as a whole? First of all, SU(2) breaking is unaffected. Secondly, a mistaken tangle–particle assignment can be

accommodated in the strand model; new forces or symmetries cannot. Therefore the strand model is not put into question.

— Could several, possibly charged, Higgs bosons exist? No such tangles seem possible – as long as a tangle with *two* Figure 96 Higgs cores in sequence is not a separate particle.
— Has some other strand model effect been overlooked? Could other elementary or composed particles exist? For example, the structure of the Higgs boson might be seen to suggest that lepton families reappear (roughly) every 125 GeV. Is that the case? The issue is not completely settled. It seems more probable that those higher tangles simply yield corrections to the Higgs mass.

In short, the existence of the standard model Higgs boson seems compatible with the strand model. The 2009 mistake about the Higgs also shows that the exploration of the strand model is not yet complete. In any case, the strand model has not been falsified by the discovery of the Higgs boson.

Page 331 Assuming that the Higgs tangle shown in Figure 96 is correct, we have an intuitive proposal for the mechanism that produces mass, namely *tail braiding*. The proposed Higgs tangle also allows a number of experimental predictions.

2012 predictions about the Higgs

— The Higgs tangle implies a Higgs boson with vanishing charge, positive parity, being elementary – as is observed.
— The Higgs tangle allows us to estimate the Higgs/Z mass ratio. Using the new, unknotted, tangle model for the W and Z bosons, the estimates are in the region of the observed values. Improving the estimates is still subject of research.
— The Higgs tangle and the strand model imply that the standard model is correct up to Planck energy, and that the Higgs mass value should reflect this. The observed Higgs mass of 125 GeV complies also with this expectation.
— Therefore, the strand model suggests that no deviations between the standard model and data should ever be observed in any experiment.
— The strand model again and consistently predicts the lack of supersymmetry.
— In the case that several Higgs bosons exist or that the braided Higgs tangle does not apply, the strand model is in trouble.
— In the case that effects, particles or interactions beyond the standard model are observed, the strand model is in trouble.

Quark-antiquark mesons

In the strand model, all three-stranded tangles apart from the leptons, as well as all four-stranded tangles represent *composite* particles. The first example are *mesons*.

In the strand model, rational tangles of three strands are quark-antiquark mesons with spin 0. The quark tangles yield a simple model of these *pseudoscalar* mesons, shown on the left-hand sides of Figure 97, Figure 99 and Figure 100. The right-hand sides of the figures show *vector* mesons, thus with spin 1, that consist of *four* strands. All tangles are rational. Inside mesons, quarks and antiquarks 'bond' at three spots that form a triangle oriented perpendicularly to the bond direction and to the paper plane. To increase clar-

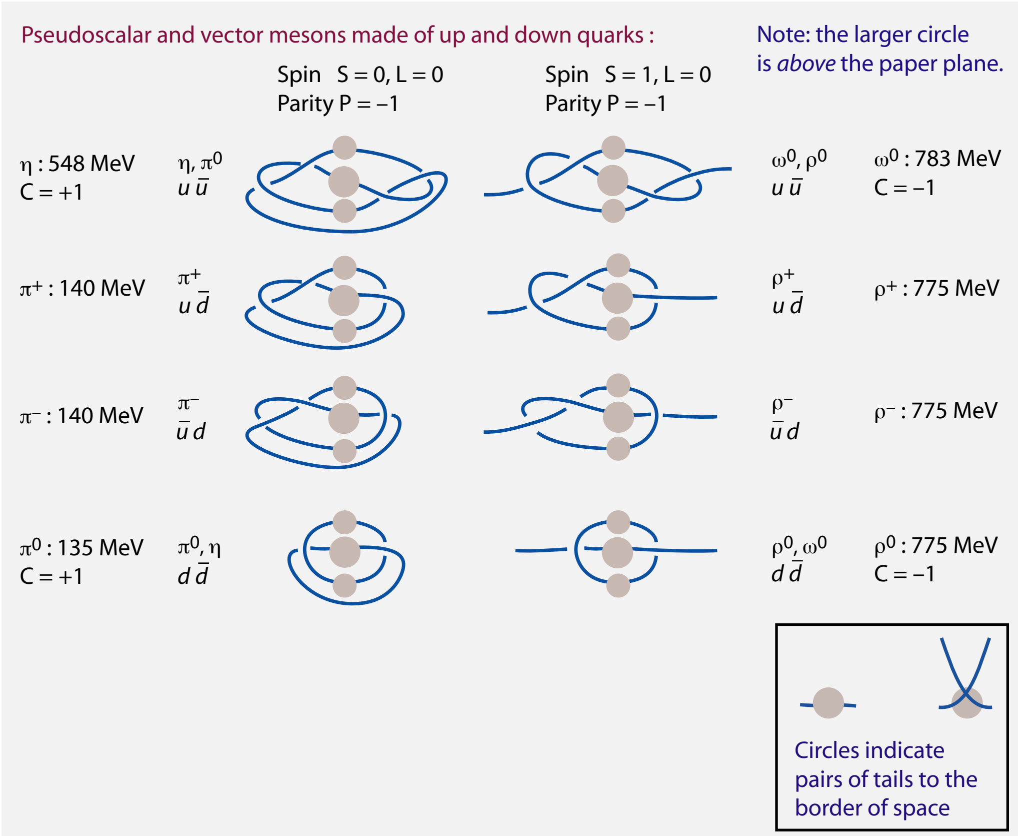


FIGURE 97 The simplest strand models for the light pseudoscalar and vector mesons (circles indicate crossed tail pairs to the border of space), with the observed mass values.

ity, the 'bonds' are drawn as circles in the figures; however, they consist of two crossed (linked) tails of the involved strands that reach the border of space, as shown in Figure 98.

With this construction, mesons made of two quarks are only possible for the type $\overline{q}\,q$. Ref. 229 Other combinations, such as $q\,q$ or $\overline{q}\,\overline{q}$, turn out to be unlinked. We note directly that Ref. 230 this model of mesons resembles the original string model of hadrons from 1973, but also Ref. 231 the Lund string model and the recent QCD string model.

To compare the meson structures with experimental data, we explore the resulting quantum numbers. As in quantum field theory, also in the strand model the parity of a particle is the product of the intrinsic parities and of wave function parity. The states with orbital angular momentum $L = 0$ are the lowest states. Experimentally, the lightest mesons have quantum numbers $J^{PC} = 0^{-+}$, and thus are pseudoscalars, or have $J^{PC} = 1^{--}$, and thus are vector mesons. The strand model reproduces these observed quantum numbers. (We note that the spin of any composite particle, such as a meson, is low-energy quantity; to determine it from the composite tangle, the tails producing the bonds – drawn as circles in the figures – must be neglected. As a result, the low-energy spin of mesons and of baryons is correctly reproduced by the strand model.)

In the strand model, the meson states are colour-neutral, or 'white', by construction, because the quark and the antiquark, in all orientations, always have opposite colours

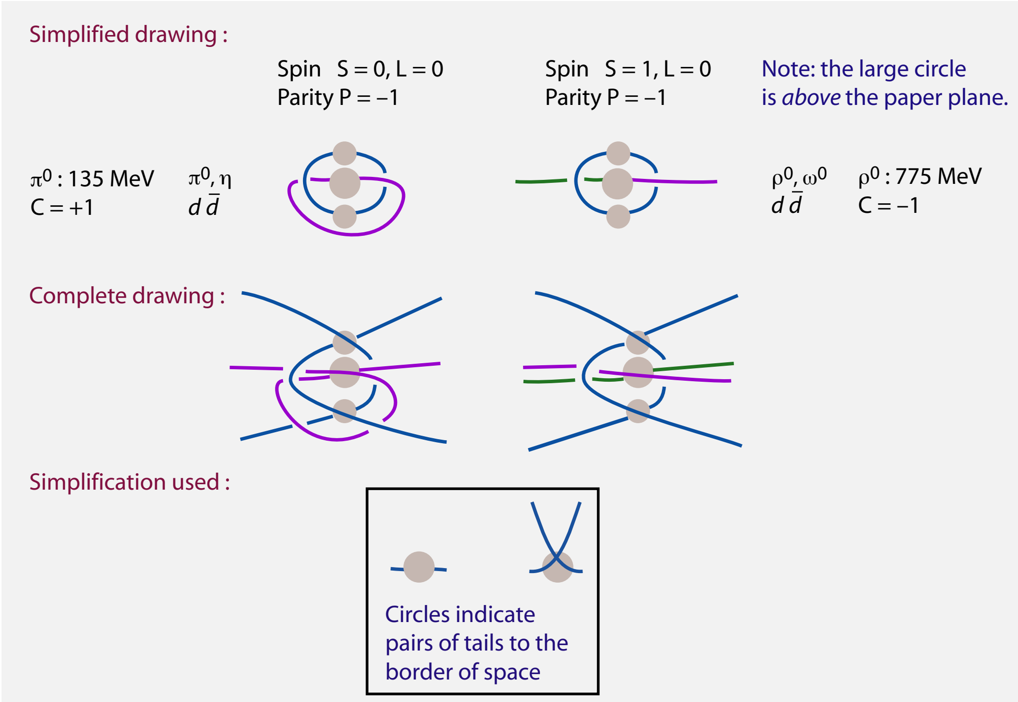


FIGURE 98 The meaning of the circles used in the tangle graphs of mesons and baryons.

that add up to white.

In the strand model, the electric charge is an integer for all mesons. Chiral tangles are charged, achiral tangles uncharged. The charge values deduced from the strand model thus reproduce the observed ones.

In experiments, no mesons with quantum numbers 0^{--}, 0^{+-}, or 1^{-+} are observed. Also this observation is reproduced by the quark tangles, as is easily checked by direct inspection. The strand model thus reproduces the very argument that once was central to the acceptance of the quark model itself.

It is important to realize that in the strand model, each meson is represented by a *tangle family* consisting of *several* tangle structures. This has three reasons. First, the 'circles' can be combined in different ways. For example, both the $u\,\bar{u}$ and the $d\,\bar{d}$ have as alternate structure a line plus a ring. This common structure is seen as the underlying reason that these two quark structures *mix*, as is indeed observed. (The same structure is also possible for $s\,\bar{s}$, and indeed, a full description of these mesons must include mixing with this state as well.) The second reason that mesons have several structures are the mentioned, more complicated braid structures possible for each quark, namely with 6, 12, etc. additional braid crossings. The third reason for additional tangle structures is the occurrence of higher-order Feynman diagrams of the weak interaction, which add yet another group of more complicated topologies that also belong to each meson.

In short, the mesons structures of Figure 97, Figure 99 and Figure 100 are only the *simplest* tangles for each meson. Nevertheless, all tangles, both the simplest and the more complicated meson tangles, reproduce spin values, parities, and all the other quantum

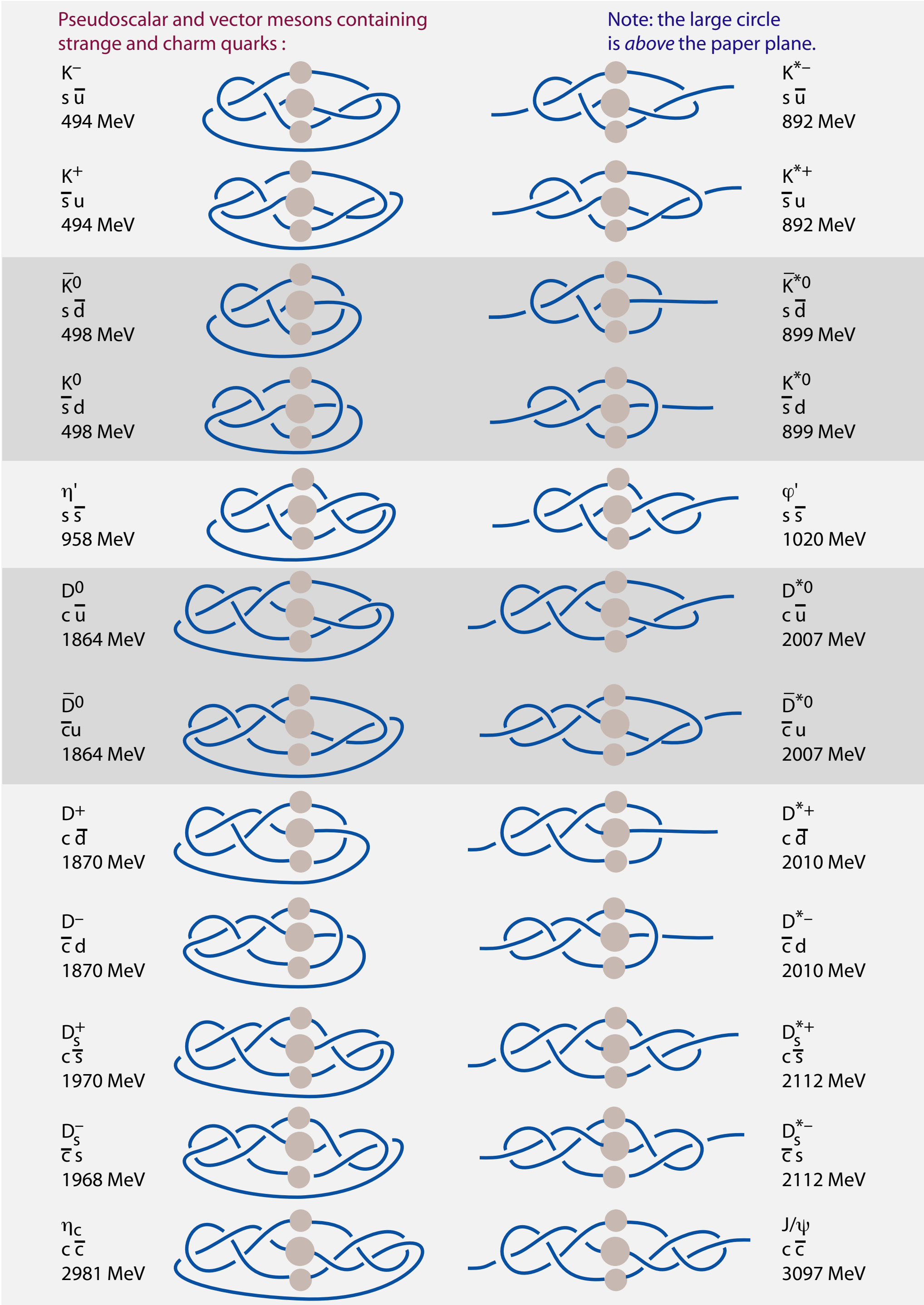


FIGURE 99 The simplest strand models for strange and charmed mesons with vanishing orbital angular momentum. Mesons on the left side have spin 0 and negative parity; mesons on the right side have spin 1 and also negative parity. Circles indicate crossed tail pairs to the border of space; grey boxes indicate tangles that mix with their antiparticles and which are thus predicted to show CP violation.

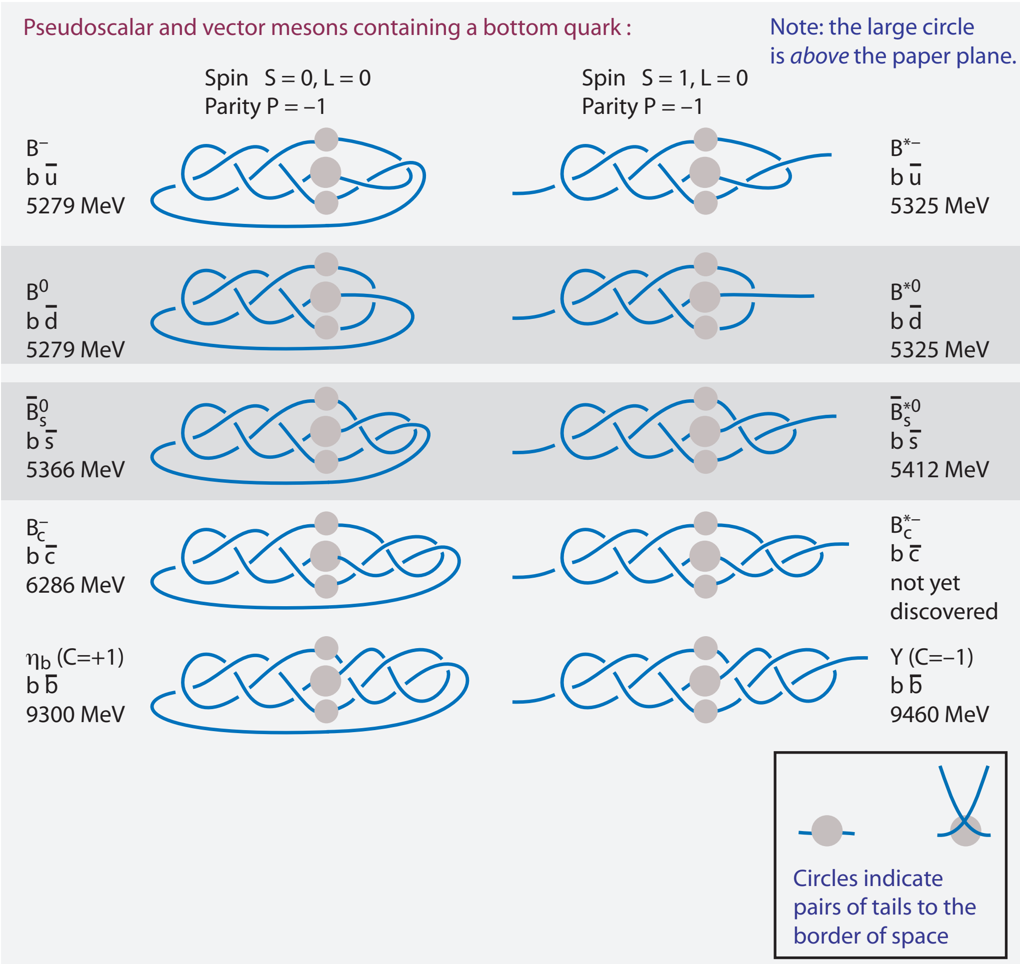


FIGURE 100 The simplest strand models for some heavy pseudoscalar and vector mesons, together with their experimental mass values. Antiparticles are not drawn; their tangles are mirrors of the particle tangles. Circles indicate crossed tail pairs to the border of space; grey boxes indicate tangles that mix with their antiparticles and which are thus predicted to show CP violation.

numbers of mesons. Indeed, in the strand model, the more complicated tangles automatically share the quantum numbers of the simplest one.

Meson form factors

The strand model also predicts directly that all mesons from Figure 97, Figure 99 and Figure 100, in fact all mesons with vanishing orbital momentum, are *prolate*. This (unsurprising) result is agreement with observations. Ref. 232 Mesons with non-vanishing orbital momentum are also predicted to be prolate. This latter prediction about meson shapes is made also by all other meson models, but has not yet been checked by experiment.

There is another way to put what we have found so far. The strand model makes the following prediction: When the meson tangles are averaged over time, the crossing densities reproduce the measured spatial, quark flavour, spin and colour part of the meson wave functions. This prediction can be checked against measured form factors and against lattice QCD calculations.

Meson masses, excited mesons and quark confinement

The strand model also allows us to understand meson masses. We recall that a *topologically complicated* tangle implies a *large* mass. With this relation, Figure 97 predicts that the π^0, η and $\pi^{+/-}$ have different masses and follow the observed meson mass sequence $m(\pi^0) < m(\pi^{+/-}) < m(\eta)$. The other mass sequences can be checked with the help of Figure 97, Figure 99 and Figure 100; there are no contradictions with observations. However, there is one limit case: the strand model predicts different masses for the ρ^0, ω, and $\rho^{+/-}$. So far, observations only partly confirm the prediction. Recent precision experiments seem to suggest that ρ^0 and $\rho^{+/-}$ have different mass; this result has not been confirmed yet. Ref. 233

More precise mass determinations will be possible with numerical calculations. This will be explored in more detail later on. Page 358 In any case, the strand model for mesons suggests that the quark masses are not so important for the determination of meson masses, whereas the details of the quark-antiquark bond are. Indeed, the light meson and baryon masses are much higher than the masses of the constituent quarks.

The relative unimportance of quark masses for many meson masses is also confirmed for the case of *excited* mesons, i.e., for mesons with orbital angular momentum L. It is well known that mesons of non-vanishing orbital angular momentum can be grouped into sets which have the same quark content, but different total angular momentum $J = L + S$. These families are observed to follow a well-known relation between total angular momentum J and mass m, called *Regge trajectories*:

$$J = \alpha_0 + \alpha_1 m^2 \tag{195}$$

Ref. 234 with an (almost) constant factor α_1 for all mesons, about 0.9 GeV/fm. These relations, the famous *Regge trajectories*, are explained in quantum chromodynamics as deriving from the linear increase with distance of the effective potential between quarks, thus from the properties of the relativistic harmonic oscillator. The linear potential itself is usually seen as a consequence of a fluxtube-like bond between quarks.

In the strand model, the fluxtube-like bond between the quarks is built-in automatically, as shown in Figure 101. All mesons have three connecting 'bonds' and these three bonds can be seen as forming one common string tube. In the simplified drawings, the bond or string tube is the region containing the circles. In orbitally excited mesons, the three bonds are expected to lengthen and thus to produce additional crossing changes, thus additional effective mass. The strand model also suggests a *linear* relation. Since the mechanism is expected to be similar for all mesons, which all have three bonding circles, the strand model predicts the *same* slope for all meson (and baryon) Regge trajectories. This is indeed observed.

In summary, the strand model reproduces meson mass sequences and quark confinement in its general properties.

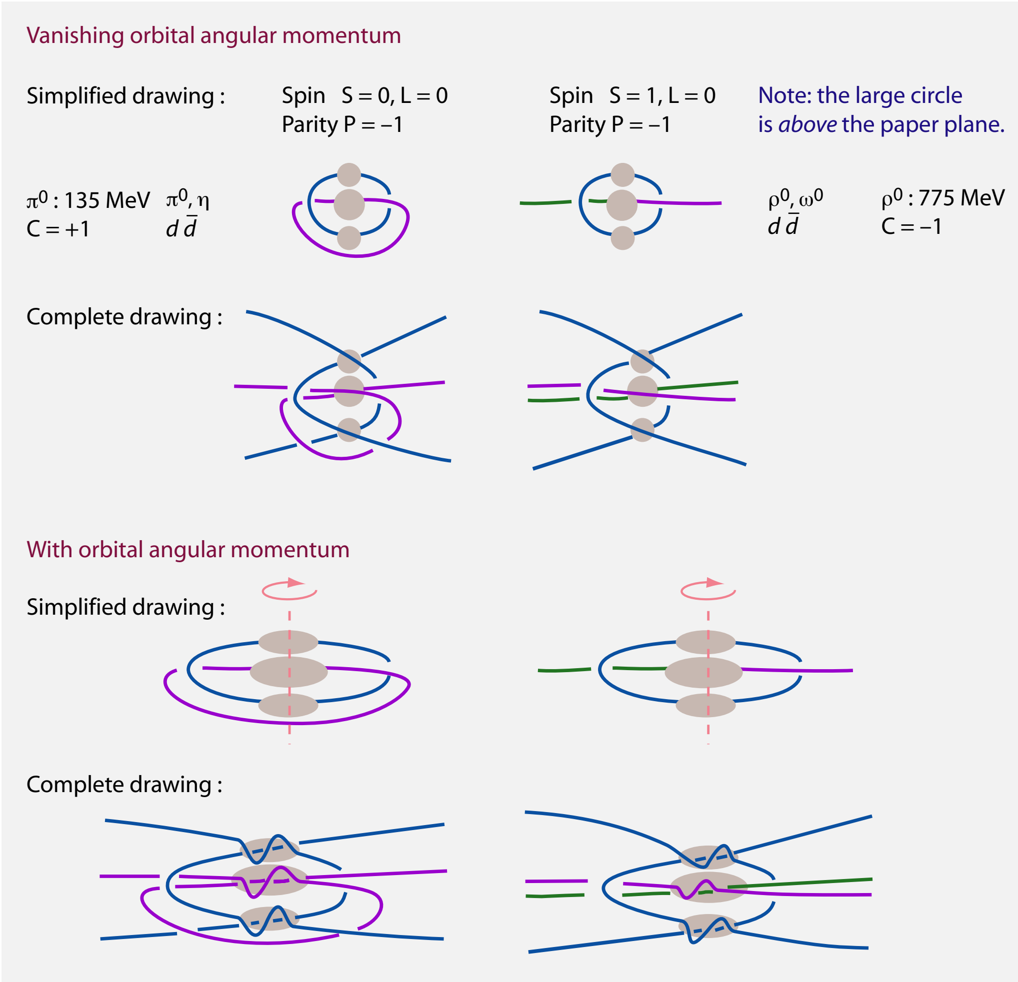


FIGURE 101 The strand model for mesons without (top) and with (bottom) orbital angular momentum.

CP violation in mesons

Ref. 235 In the weak interaction, the product CP of C and P parity is usually conserved. However, rare exceptions are observed for the decay of the K^0 meson and in various processes that involve the B^0 and B^0_s mesons. In each of these exceptions, the meson is found to Ref. 233 mix with its own antiparticle. CP violation is essential to explain the matter–antimatter asymmetry of the universe.

The strand model allows us to deduce whether the mixing of a meson with its own antiparticle is possible or not. As expected, only neutral mesons are candidates for such mixing, because of charge conservation. In the strand model, particle–antiparticle mixing is possible whenever the transition from a neutral meson to its antiparticle is possible in *two* ways: by taking the mirror of the meson tangle or by shifting the position of the binding strands. All mesons for which this is possible are shown in grey boxes in Figure 97, Figure 99 and Figure 100. The strand model also makes it clear that such mixing requires shifting of the bonds; this is a low-probability process that is due to the weak interaction. The strand model thus predicts that the weak interaction violates CP invariance in mesons that mix with their antiparticles.

Since the spin 1 mesons decay strongly and thus do not live long enough, the small effect of CP violation is de facto only observed in pseudoscalar, spin-0 mesons. The strand model thus predicts observable mixings and CP violation for the mesons pairs $K^0 - \overline{K}^0$, $D^0 - \overline{D}^0$, $B^0 - \overline{B}^0$, $B^0_s - \overline{B}^0_s$. The prediction by the strand model corresponds Ref. 233 precisely to those systems for which CP violation is actually observed. (CP violation in D mesons was finally discovered at CERN in 2011, after it was predicted both by the standard model and the strand model, in earlier editions of this volume.)

In the strand model, meson–antimeson mixing is possible because the various quarks are braided strands. Because of this braid structure, the existence of meson–antimeson mixing is a consequence of the existence of three quark generations. The meson structures also make it clear that such mixings would not be possible if there were no third quark generation. The strand model thus reproduces the usual explanation of CP violation as the result of three quark generations.

For the strong and the electromagnetic interaction, the strand model predicts that there is no mixing and no CP violation, because gluons and photons do not change particle topology. Therefore, the strand model suggests the absence of axions. The lack of Page 353 a suitable tangle for axions, shown later on, then turns this suggestions into a prediction.

In summary, the existence of CP violation in the weak interactions and the lack of CP violation in the strong interaction are natural consequences of the strand model.

Other three-stranded tangles

In the strand model, the omitted complicated tangles made of three strands are either higher-order propagating versions of the tangles just presented or composites of one-stranded or two-stranded particles.

The three-strand analog of the graviton – three parallel, but twisted strands – is not Challenge 195 s an elementary particle, but a composed structure.

Spin and three-stranded particles

Why do three strands sometimes form a spin 0 particle, such as the elementary Higgs boson, sometimes a spin 1/2 particle, such as the elementary electron, and sometimes a spin 1 particle, such as a composed meson? The answer depends on how the strands are free to move against each other.

The Higgs tangle appears through tangling of vacuum strands, and inherits the zero spin of vacuum. The W and Z tangles have a special property: two strands can rotate around the third; this makes them bosons as well, but of spin 1. Fermion tangles have neither property; their core can only rotate through the belt trick; thus they are fermions.

Summary on three-stranded tangles

A number of elementary particles are made of trhee strands: The massive W and Z, the gluons, the leptons and the Higgs boson. Their tangles reproduce all their observed quantum numbers. The tangles also imply that neutrinos and anti-neutrinos differ, are massive, and are Dirac particles.

Page 330 The strand model (corrected in 2012) also predicts that, apart from the mentioned particles, no other elementary particle made of three strands exist in nature.

In the case of *composite* particles made of three strands, the strand model proposes tangles for all pseudoscalar mesons; the resulting quantum numbers and mass sequences match the observed values.

TANGLES OF FOUR AND MORE STRANDS

If we add one or more strand to a three-strand tangle, no additional class of tangles appears. The tangle classes remain the same as in the three-strand case. In other words, *no additional elementary particles* arise in the strand model. To show this, we start our exploration with the *rational* tangles.

We saw above that the rational tangles made of four strands represent the vector mesons. We have already explored them together with the scalar mesons. But certain more complicated rational tangles are also important in nature, as we consist of them.

Baryons

In the strand model, rational tangles made of five or six strands are baryons. The quark tangles of the strand model yield the tangles for baryons in a natural way, as Figure 102 shows. Again, not all quark combinations are possible. First of all, quark tangles do not allow mixed $q\,q\,\overline{q}$ or $q\,\overline{q}\,\overline{q}$ structures, but only $q\,q\,q$ or $\overline{q}\,\overline{q}\,\overline{q}$ structures. In addition, the tangles do not allow (fully symmetric) spin 1/2 states for $u\,u\,u$ or $d\,d\,d$, but only spin 3/2 states. The model also naturally predicts that there are only two spin 1/2 baryons made of u and d quarks. All this corresponds to observation. The tangles for the simplest baryons are shown in Figure 102.

The electric charges of the baryons are reproduced. In particular, the tangle topologies imply that the proton has the same charge as the positron. Neutral baryons have topologically achiral structures; nevertheless, the neutron differs from its antiparticle, as can be deduced from Figure 102, through its three-dimensional shape. The Δ baryons
Page 390 have different electric charges, depending on their writhe.

Baryons are naturally colour-neutral, as observed. The model also shows that the baryon wave function usually cannot be factorized into a spin and quark part: the nucleons need *two* graphs to describe them, and tangle shapes play a role. Baryon parities are reproduced; the neutron and the antineutron differ. All this corresponds to known baryon behaviour. Also the observed baryon shapes (in other words, the baryon quadrupole
Ref. 232 moments) are reproduced by the tangle model.

The particle masses of proton and neutron differ, because their topologies differ. However, the topological difference is 'small', as seen in Figure 102, so the mass difference is small. The topological difference between the various Δ baryons is even smaller, and indeed, their mass difference is barely discernible in experiments.

The strand model naturally yields the baryon octet and decuplet, as shown in Figure 103 and Figure 104. In general, complicated baryon tangles have higher mass than simpler ones, as shown in the figures; this is also the case for the baryons, not illustrated here, that include other quarks. And like for mesons, baryon Regge trajectories are due

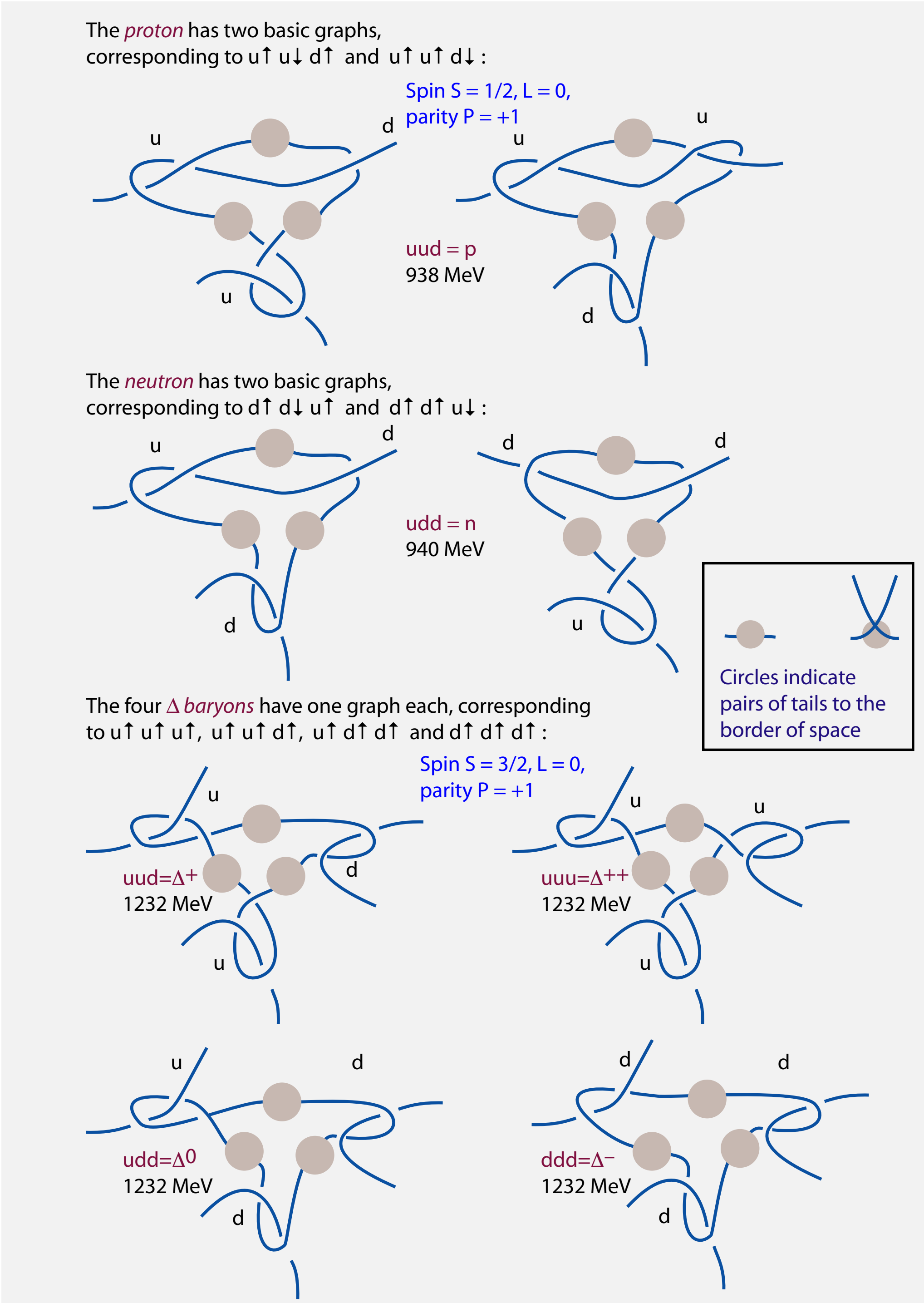


FIGURE 102 The simplest strand models for the lightest baryons made of up and down quarks (circles indicate linked tail pairs to the border of space), together with the measured mass values.

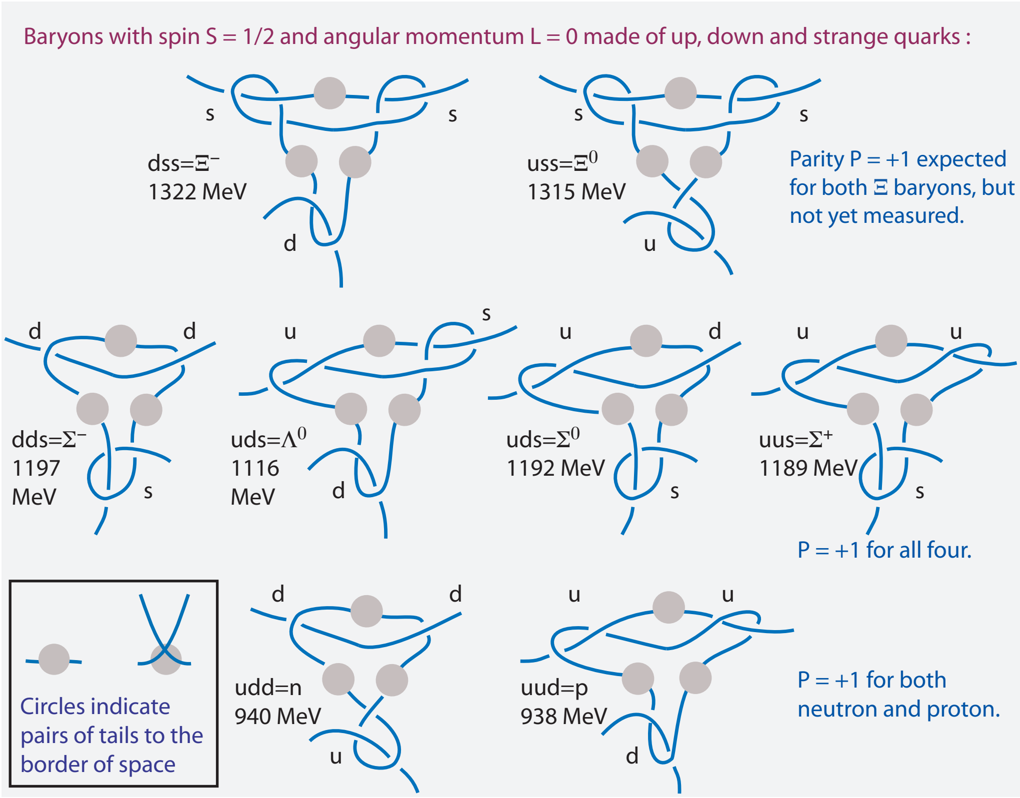


FIGURE 103 One tangle (only) for each baryon in the lowest J=L+S=1/2 baryon octet (circles indicate linked tail pairs to the border of space), together with the measured mass values.

to 'stretching' and tangling of the binding strands. Since the bonds to each quark are again (at most) three, the model qualitatively reproduces the observation that the Regge slope for all baryons is the same and is equal to that for mesons. We note that this also implies that the quark masses play only a minor role in the generation of hadron masses; this old result from QCD is thus reproduced by the strand model.

The arguments presented so far only reproduce mass sequences, not mass values. Actual hadron mass calculations are possible with the strand model: it is necessary to compute the number of crossing changes each tangle produces. There is a chance, but no certainty, that such calculations might be simpler to implement than those of lattice QCD.

Tetraquarks and exotic mesons

Ref. 220 Among the exotic mesons, tetraquarks are the most explored cases. It is now widely be-
Ref. 236 lieved that the low-mass scalar mesons are tetraquarks. In the strand model, tetraquarks are possible; an example is given in Figure 105. This is a six-stranded rational tangle. Spin, parities and mass sequences from the strand model seem to agree with observations. If the arrangement of Figure 105 would turn out to be typical, the tetraquark looks more like a bound pair of two mesons and not like a state in which all four quarks are bound in equal way to each other. On the other hand, a tetrahedral arrangement of quarks might

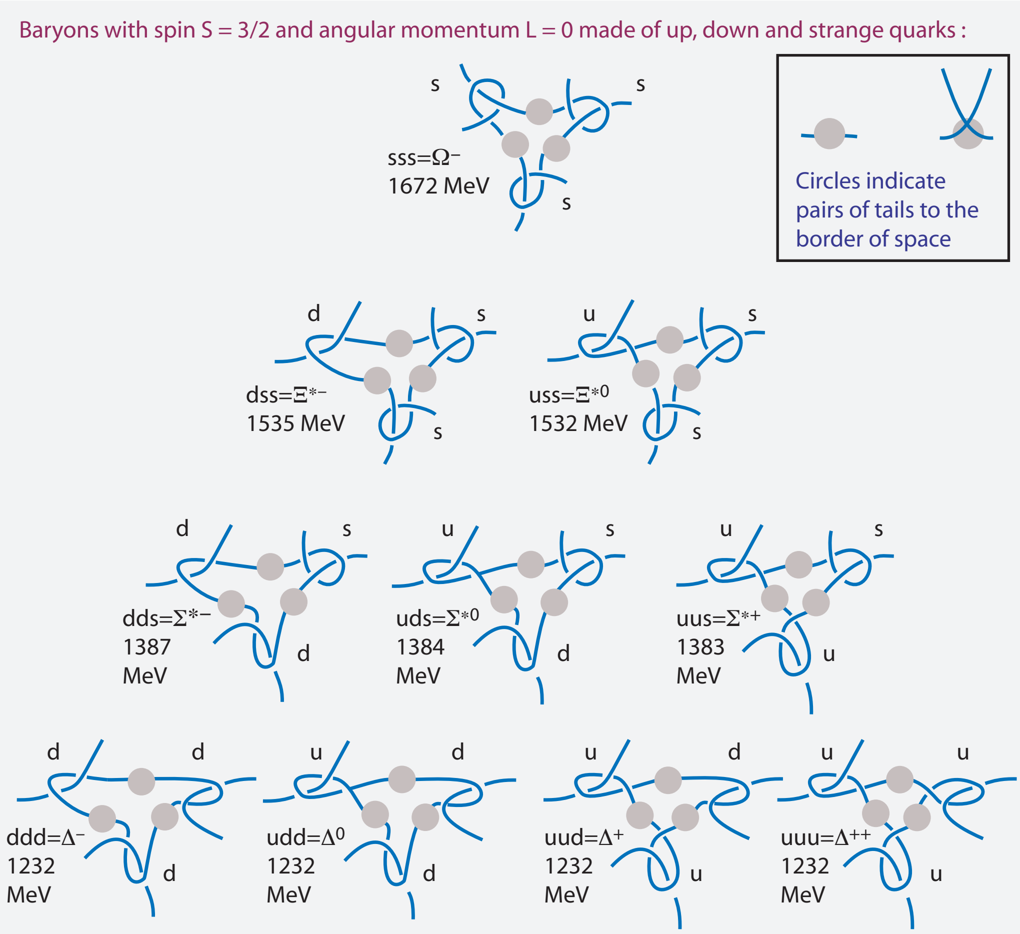


FIGURE 104 One tangle for each baryon in the lowest J=3/2 baryon decuplet (circles indicate linked tail pairs to the border of space), together with the measured mass values.

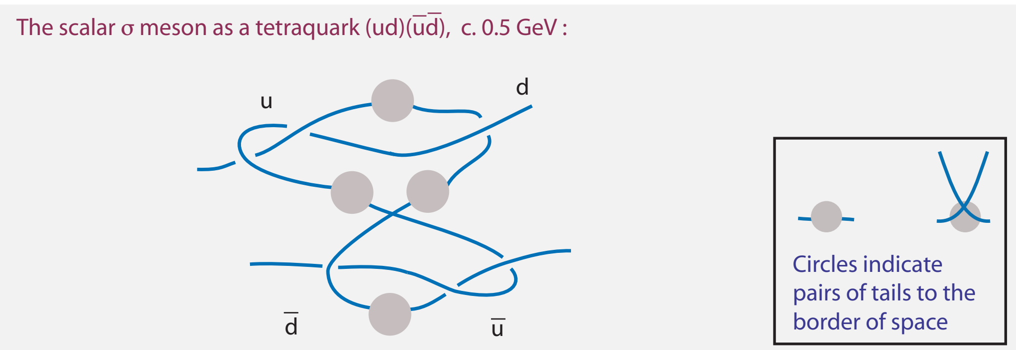


FIGURE 105 The strand model for a specific tetraquark (circles indicate linked tail pairs to the border of space).

also be possible. The details of this topic are left for future exploration.

The strand model makes an additional statement: knotted (hadronic) strings in quark–antiquark states are impossible. Such states have been proposed by Niemi. In the strand model, such states would not be separate mesons, but usual mesons with one or several added virtual weak vector bosons. This type of exotic mesons is therefore predicted not to exist.

Ref. 222

The situation for glueballs, which are another type of exotic mesons, has already been discussed above.

Page 344, page 339

Other tangles made of four or more strands

We do not need to explore other *prime* tangles or *locally knotted* tangles made of four or more strands. They are either not allowed or are higher-order versions of rational tangles, as explained already in the case of two and three strands. We also do not need to explore *separable* tangles. Separable tangles are composite of tangles with fewer strands.

One class of tangles remains to be discussed: *braided* tangles of four or more strands. Now, a higher-order perturbation of the weak interaction can always lead to the topological entanglement of some vacuum strand with a tangle of fewer strands. Braided tangles of four or more strands are thus higher-order propagating states of three-stranded leptons or hadrons.

We can also state this in another way. There are no tangles of four or more strands that are more tangled than the trivial tangle but less tangled than the lepton tangles. Therefore, no additional elementary particles are possible. In short, *the tangle model does not allow elementary particles with four or more strands.*

Glueballs

Ref. 219, Ref. 220

There is no observational evidence for glueballs yet, even though simulations of QCD on the lattice predict the existence of several such states in the 1.5 GeV/c^2 mass range. The lack of experimental confirmation is usually explained by the strong background noise in the reaction that produces glueballs, and by the expected strong mixing with mesons of similar quantum numbers. The experimental search for glueballs is still ongoing.

The often conjectured glueball could be made of two or three gluons. (The lowest-mass glueball is usually expected to be made of two gluons.) In the strand model, such structures would be tangles made of six or nine strands.

However, the masslessness of gluons and their spin do not seem to allow such a tangle. The argument is not watertight, however, and the issue is still subject of research.

Ref. 221

Whatever the situation for glueballs might be, the strand model of gluons seems in contrast with the models of glueballs as knots that were proposed by Buniy and Kephart or by Niemi. These models are based on *closed* knots, not on tangles with tails. The strand model does not seem to allow real particles *of zero spin* that are composed of gluons. On the other hand, if closed knots were somehow possible in the strand model, they would imply the existence of glueballs.

Ref. 222

In summary, the issue of glueballs is not settled; a definitive solution might even lead to additional checks of the strand model.

The mass gap problem and the Clay Mathematics Institute

Ref. 223 The Clay Mathematics Institute offers a large prize to anybody who proves the following statement: *For any compact simple non-Abelian gauge group, quantum gauge theory exists in continuous, four-dimensional space-time and produces a mass gap.* This is one of their so-called *millennium problems*.

The strand model does not allow arbitrary gauge groups in quantum field theory. According to the strand model, the only compact simple non-Abelian gauge group of interest is SU(3), the gauge group of the strong nuclear interaction. And since the strand model does not seem to allow for glueballs, for SU(3) an effective mass gap of the order of the Planck mass is predicted. (If glueballs would exist in the strand model, the mass gap would still exist but be smaller.) Indeed, the strand model explains the short range of the strong interaction as a consequence of the details of Reidemeister III moves and the quark tangle topology.

The strand model further states that space-time and gauge groups are low-energy approximations that arise through time-average of strands and their crossings. Neither points nor fields exist at a fundamental level; points and fields are approximations to strands. According to the strand model, the *quantum* properties of nature result from the extension of strands. As a consequence, the strand model denies the existence of *any quantum* gauge theory as a separate, exact theory on *continuous* space-time.

The strand model does predict a mass gap for SU(3); but the strand model also denies the existence of quantum gauge theory for any other compact simple non-Abelian gauge group. And even in the case of SU(3) it denies – like for any other gauge groups – the existence of a quantum gauge theory on continuous space-time. As deduced above, the strand model allows only the three known gauge groups, and allows their existence only in the non-continuous strand model of space-time. In short, it is *impossible* to realize the wish of the Clay Mathematics Institute.

Page 276

Summary on tangles made of four or more strands

By exploring all possible tangle classes in detail, we have shown that *every* localized structure made of strands has an interpretation in the strand model. In particular, the strand model makes a simple statement on any tangle made of four or more strands: such a tangle is *composite* of the elementary tangles made of one, two or three strands. In other terms, there are *no* elementary particles made of four or more strands in nature.

The strand model states that each possible tangle represents a physical particle system: an overview is given in Table 13. The mapping between tangles and particles is only possible because (infinitely) many tangles are assigned to each massive elementary particle.

The result of this exploration is that the strand model limits the number of elementary particles to those contained in the standard model of particle physics.

TABLE 13 The match between tangles and particles in the strand model.

Strands	Tangle	Particle	Type
1	unknotted	elementary	vacuum, (unbroken) gauge boson
1	knotted	–	not in the strand model
2	unknotted	composed	composed of simpler tangles
2	rational	elementary	quark or graviton
2	prime, knotted	–	not in the strand model
3	unknotted	composed	composed of simpler tangles
3	braided	elementary	lepton
3	rational	elementary or composed	leptons
3	prime, knotted	–	not in the strand model
4 & more	like for 3 strands	all composed	composed of simpler tangles

FUN CHALLENGES AND CURIOSITIES ABOUT PARTICLE TANGLES

In the strand model, mass appears due to tail braiding. But mass is also due to tangle rotation and fluctuation. How do the two definitions come together? Challenge 196 s

* *

The following statement seems absurd, but is correct:

▷ The tangle model implies that all elementary particles are point-like, *without* internal structure.

Indeed, if at all, the strand model implies deviations from point-like behaviour only at Planck scale; particles are point-like for all practical purposes.

* *

In the strand model, only crossing switches are observable. How then can the specific tangle structure of a particle have any observable effects? In particular, how can quantum numbers be related to tangle structure, if the only observables are due to crossing changes? Challenge 197 e

* *

No neutral weak currents that change strangeness or other flavours are observed. In the strand model this observation is a consequence of the tangle shape of the Z boson.

* *

Ref. 237 In 2014, Marek Karliner predicted the existence of six-quark states. Can the strand model reproduce them? Can it settle whether they are molecules of three mesons or genuine Challenge 198 r six-quark states?

* *

Challenge 199 e Can you use the strand model to show that pentaquarks do not exist?

* *

Ref. 238 What is the relation of the model shown here to the ideas of Viro and Viro on skew lines?

* *

The most prominent proponent of the idea that particles might be knots was, in 1868, William Thomson–Kelvin. He proposed the idea that different atoms might be differently 'knotted vortices' in the 'ether'. The proposal was ignored – and rightly so – because it did not explain anything: neither the properties nor the interactions of atoms were explained. The proposal simply had no relation to reality. In retrospect, the main reason for this failure was that elementary particles and quantum theory were unknown at the time. (Ref. 239)

* *

Purely topological models for elementary particles have been proposed and explored by various scholars in the past. But only a few researchers ever proposed specific topological structures for each elementary particle. Such proposals are easily criticized, so that it is easy to make a fool of oneself; any such proposal thus needs a certain amount of courage.

— Herbert Jehle modelled elementary particles as closed knots already in the 1970s. However, his model did not reproduce quantum theory, nor does it reproduce all particles known today. (Ref. 240)
— Ng Sze Kui has modelled mesons as knots. There is however, no model for quarks, leptons or bosons, nor a description for the gauge interactions. (Ref. 143)
— Tom Mongan has modelled elementary particles as made of three strands that each carry electric charge. However, there is no connection with quantum field theory or general relativity. (Ref. 241)
— Jack Avrin has modelled hadrons and leptons as Moebius bands, and interactions as cut-and-glue processes. The model however, does not explain the masses of the particles or the coupling constants. (Ref. 139)
— Robert Finkelstein has modelled fermions as knots. This approach, however, does not explain the gauge properties of the interactions, nor most properties of elementary particles. (Ref. 141)
— Sundance Bilson-Thompson, later together with his coworkers, modelled elementary fermions and bosons as structures of triple ribbons. The leather trick is used, like in the strand model, to explain the three generations of quarks and leptons. This is by far the most complete model from this list. However, the origin of particle mass, of particle mixing and, most of all, of the gauge interactions is not explained. (Ref. 140)

* *

Strands are *not* superstrings. In contrast to superstrings, strands have a fundamental principle. (This is the biggest conceptual difference.) The fundamental principle for strands is not fulfilled by superstrings. In contrast to superstrings, strands have no tension, no supersymmetry and no own Lagrangian. (This is the biggest physical difference.) Because strands have no tension, they cannot oscillate. Because strands have no supersymmetry, general relativity follows directly. Because strands have no own Lagrangian, particles are tangles, not oscillating superstrings, and quantum theory follows directly. In fact, the definitions of particles, wave functions, fields, vacuum, mass and horizons differ completely in the two approaches.

In contrast to superstrings, strands describe the number of gauge interactions and of particle generations. In contrast to superstrings, strands describe quarks, hadrons, confinement, Regge behaviour, asymptotic freedom, particle masses, particle mixing and coupling constants. In the strand model, in contrast to 'open superstrings', no important configuration has ends. In contrast to open or closed superstrings, strands move in three spatial dimensions, not in nine or ten; strands resolve the anomaly issue without higher dimensions or supersymmetry, because unitarity is automatically maintained, by construction; strands are not related to membranes or supermembranes. In the strand model, no strand is 'bosonic' or 'heterotic', there is no E(8) or SO(32) gauge group, there are no general 'pants diagrams' for all gauge interactions, there is no AdS/CFT duality, there is no 'landscape' with numerous vacuum states, and there is no 'multiverse'. In contrast to superstrings, strands are based on Planck units. And in contrast to superstrings, strands yield the standard model of elementary particles without any alternative.
Ref. 144 In fact, not a single statement about superstrings is applicable to strands.

* *

Strands do not require higher dimensions. On the other hand, it can be argued that strands do produce an additional non-commutative structure at each point in space. In a sense, when strands are averaged over time, a non-commutative inner space is created at each point in space. As a result, when we focus at a specific spatial position over somewhat longer times scales than the Planck time, we can argue that, at that point of space, nature is described by a product of three-dimensional space with an internal, non-commutative space. Since many years, Alain Connes and his colleagues have explored
Ref. 242 such product spaces in detail. They have discovered that with an appropriately chosen non-commutative inner space, it is possible to reproduce many, but not all, aspects of the standard model of particle physics. Among others, choosing a suitable non-commutative space, they can reproduce the three gauge interactions; on the other hand, they cannot reproduce the three particle generations.

Connes' approach and the strand model do not agree completely. One way to describe the differences is to focus on the relation of the inner spaces at different points of space. Connes' approach assumes that each point has its own inner space, and that these spaces are not related. The strand model, instead, implies that the inner spaces of neighbouring points are related; they are related by the specific topology and entanglement of the involved strands. For this very reason the strand model does allow to understand the origin of the three particle generations and the details of the particle spectrum.

There are further differences between the two approaches. Connes' approach assumes that quantum theory and general relativity, in particular, the Hilbert space and the spatial manifold, are given from the outset. The strand model, instead, deduces these structures from the fundamental principle. And, as just mentioned, Connes' approach is not

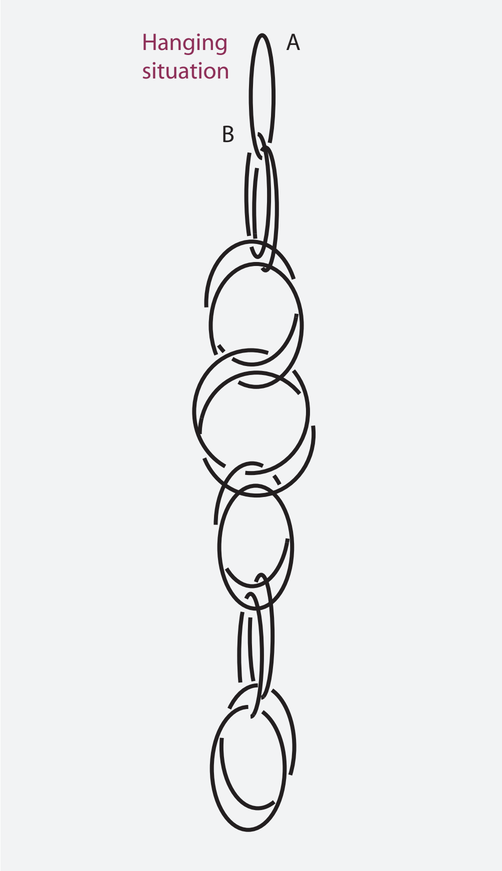


Figure 106 A ring chain gives an impression of motion along the chain, when holding ring B while dropping ring A.

unique or complete, whereas the strand model seems to be. Of the two, only the strand model seems to be unmodifiable, or 'hard to vary'.

* *

The strand model implies that there is *nothing new* at small distances. At small distances, or high energies, nature consists only of strands. Thus there are no new phenomena there. Quantum theory states that at small scales, nothing new appears: at small scales, there are *no* new degrees of freedom. For example, quantum theory states that there is no kingdom *Lilliput* in nature. The strand model thus confirms the essence of quantum theory. And indeed, the strand model predicts that between the energy scale of the heaviest elementary particle, the top quark, 173 GeV, and the Planck energy, 10^{19} GeV, nothing is to be found. There is a so-called *energy desert* – empty of interesting features, particles or phenomena – in nature.

* *

Most ropes used in sailing, climbing or other domains of everyday life are produced by braiding. Searching for 'braiding machine' on the internet yields a large amount of videos. Searching for 'LEGO braiding machine' shows the most simple and beautiful

FIGURE 107 The ring chain trick produces an illusion of motion (mp4 film © Franz Aichinger). Can more rings be added in horizontal directions?

Challenge 200 e

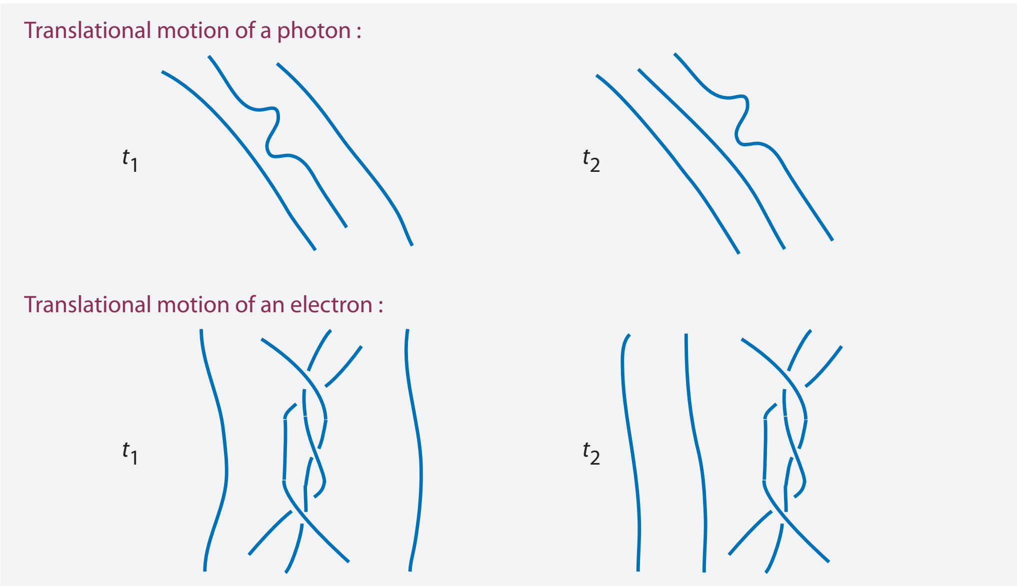


FIGURE 108 Motion of photons and electrons through strand hopping.

examples and allows you to see how they work.

* *

Not all tangle assignments are self-evident at first sight. Figure 109 shows a tangle whose status in the strand model is not clear. Can you explain what the tangle represents?

Challenge 201 e

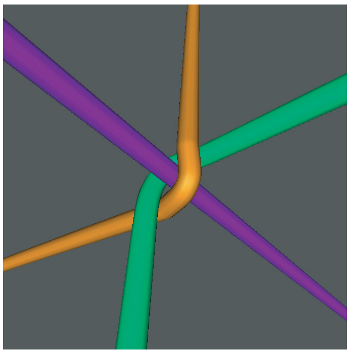


FIGURE 109 A discarded candidate tangle for the W boson.

* *

Page 153
Challenge 202 ny
What is the effect of shivering on braiding, and thus on weak particle mixing, on particle tangle families and on the number of generations?

* *

Are all bosons made of strands whose ends are exactly opposite to each other at spatial infinity? Photon, graviton, gluon, W, Z and the virtual Higgs comply. The unbroken ones are axial, the broken ones are flat. Is there a reason or a sense for this issue?
Challenge 203 e

CPT invariance

CPT invariance is a fundamental property of quantum field theory. In the strand model, charge conjugation C is modelled as a mirror transformation of the tangle; parity P is modelled as the change of sign of the belt trick of the tangle core; and motion inversion T is modelled as the inverse motion of the core of a particle tangle.

In other words, CPT invariance is natural in the strand model. Therefore, the strand model predicts that particles and antiparticles have the same g-factor, the same dipole moment, the same mass, the same spin, exactly opposite charge value, etc. All this is also predicted by quantum field theory, and is confirmed by experiment.

Motion through the vacuum – and the speed of light

Up to now, one problem was left open: How can a particle, being a tangle of infinite extension, move through the web of strands that makes up the vacuum? An old trick, known already in France in the nineteenth century, can help preparing for the idea of particle motion in space. Figure 106 shows a special chain that is most easily made with a few dozen key rings. If the ring B is grabbed and the ring A released, this latter ring seems to fall down along the whole chain in a helical path, as shown in the film of Figure 200. If you have never seen the trick, try it yourself; the effect is astonishing. In reality, this is an optical illusion. No ring is actually falling, but the sequence of rings moves in a way that creates the impression of ring motion. And this old trick helps us to solve a number of issues about particle motion that we swept under the carpet so far.
Ref. 243

The main idea on particle motion in the strand model is the following:

▷ ***Translational particle motion*** is also due to strand substitution, or 'strand hopping'.

A schematic illustration of translational motion is given in Figure 108. In the strand model, contrary to the impression given so far, a tangle does not always need to move as a whole along the strand. This is seen most easily in the case of a photon. It is easy to picture that the tangle structure corresponding to a photon can also hop from strand to strand. At any stage, the structure is a photon; but the involved strand is never the same.

The idea of motion through strand hopping also works for massive particles. The motion of a massive particle, such as an electron, is shown schematically in Figure 108. The figure shows that through a tail unbraiding, the structure that describes an electron can get rid of one strand and grab a new one. This process has a low probability, of course. In the strand model, this is one reason that massive particles move more slowly than light, even if the first approximation yields a zero mass value.

We note that this explanation of motion is important also for the mapping from strand diagrams to Feynman diagrams. For many such diagrams, for example for the annihilation of particles and antiparticles in QED, strand hopping and tail unbraiding play a role. Without them, the mapping from strands to quantum field theory would not be possible.

In summary, tangles of massive particles *can* move through the vacuum using hopping – via tail unbraiding – and this naturally happens more slowly than the motion of photons, which do not need any process at the border of space to hop. The speed of photons is thus a limit speed for massive particles; special relativity is thus recovered.

> "The ground of science was littered with the corpses of dead unified theories."
>
> Freeman Dyson

SUMMARY ON MILLENNIUM ISSUES ABOUT PARTICLES AND THE VACUUM

We have discovered that the strand model makes a strong statement: elementary particles can only be made of one, two or three tangled strands. Each massive elementary particle is represented by an infinite *family* of rational tangles of fixed strand number. The family members differ by the added number of braids made of three strands with 6 crossings, i.e., by the different numbers of virtual Higgs bosons.

For *one-stranded* particles, the strand model shows that the photon is the only possibility. For *two-stranded* particles, the strand model shows that there are precisely three generations of two massive quarks. For *three-stranded* elementary particles, the strand model shows that there is a Higgs boson, 8 gluons, the W, the Z, and three generations of leptons. Neutrinos and antineutrinos differ and are massive Dirac particles. The strand model thus predicts that the neutrino-less double-beta decay will *not* be observed. Glueballs most probably do not exist.

The strand model uses the tangle assignments of Figure 110 and Figure 111 to explain the origin of all quantum numbers of the observed elementary particles. The strand model reproduces the quark model, including all the allowed and all the forbidden hadron states. For mesons and baryons, the strand model predicts the correct mass sequences and quantum numbers. Therefore, we have also completed the argument that

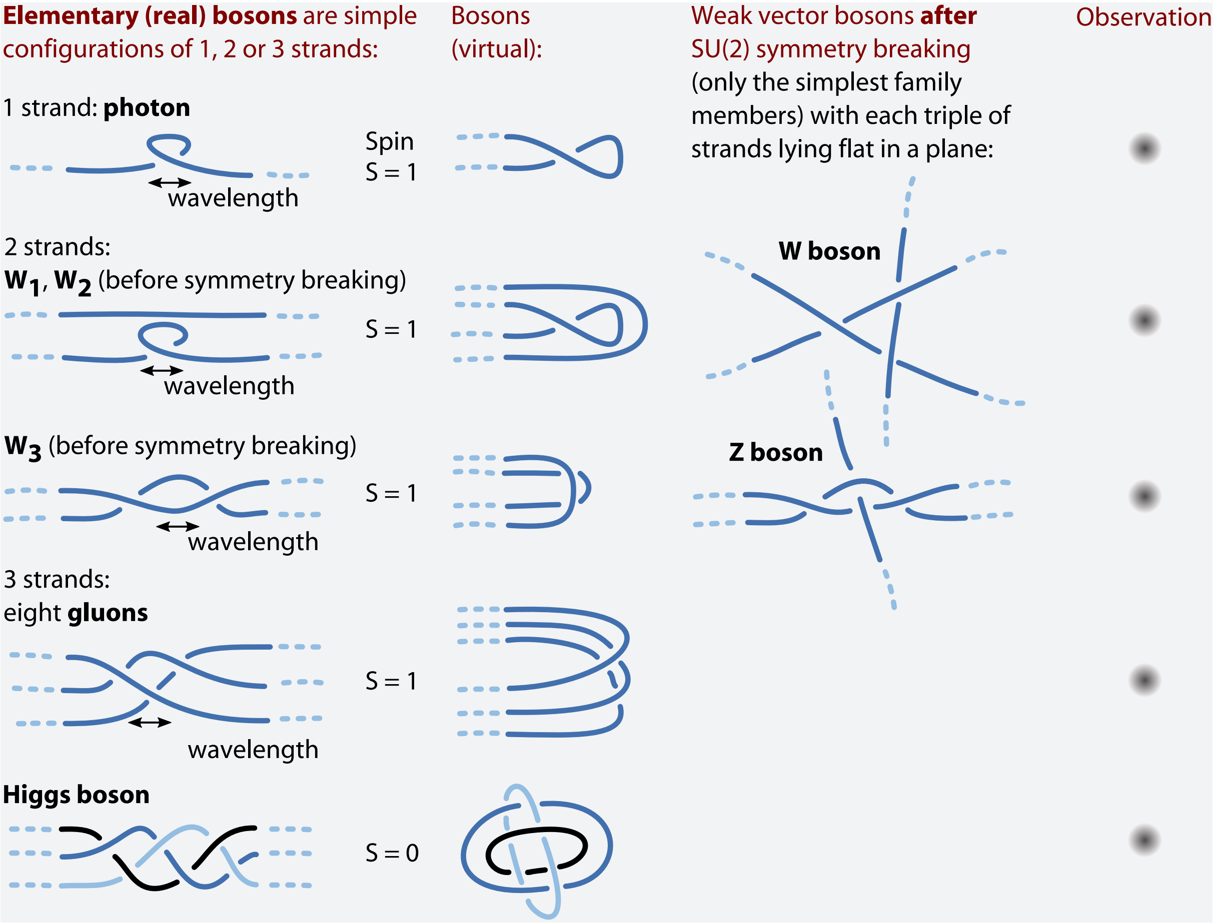


FIGURE 110 The tangle models for the elementary bosons. These tangles determine the spin values, the corresponding propagators, and ensure that the massless photons and gluons move with the speed of light. No additional elementary bosons appear to be possible.

Page 155 all observables in nature are due to crossing switches. Tetraquarks are predicted to exist. A way to calculate hadron form factors is proposed.

In the strand model, all tangles are mapped to known particles. The strand model predicts that *no* elementary particles outside the standard model exist, because no tangles are left over. For example, there are no axions, no leptoquarks and no supersymmetric particles in nature. The strand model also predicts the lack of other gauge bosons and other interactions. In particular, the strand model – corrected in 2012 – reproduces the Page 330 existence of exactly one Higgs boson. In fact, any new elementary particle found in the future would contradict and invalidate the strand model.

In simple words, the strand model explains why the known elementary particles exist and why others do not. We have thus settled two further items from the millennium list Page 164 of open issues. In fact, the deduction of the elementary particle spectrum given here is, the first and, at present, also the *only* such deduction in the research literature.

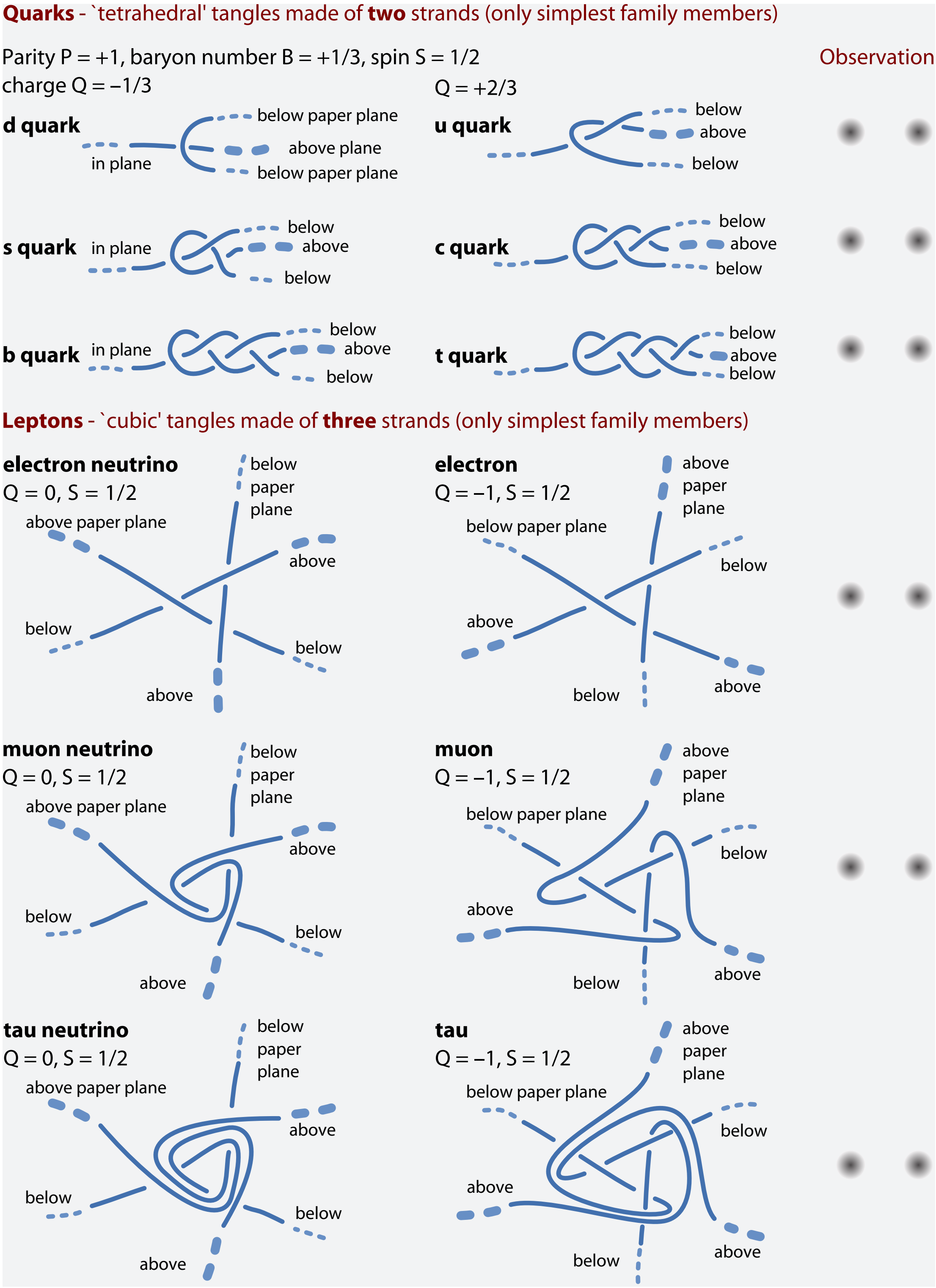


FIGURE 111 Elementary fermions are described by rational, i.e., unknotted tangles. Their structures lead to coupling to the Higgs, as illustrated in Figure 130, produce positive mass values, and limit the number of generations to 3. The tangles determine the specific fermion propagators. The tethers of the quark tangles follow the axes of a tetrahedron. The neutrino cores are simpler when seen in three dimensions: they are simply twisted triples of strands. The tethers of all lepton tangles approach the three coordinate axes at large distances from the core. No additional elementary fermions appear to be possible.

The omnipresent number 3

The strand model shows that the number 3 that appears so regularly in the standard model of particle physics – 3 generations, 3 interactions, charge values $e/3$ and $2e/3$ of quarks (as shown below), 3 colours and SU(3) – is, in each case, a consequence of the three-dimensionality of space. In fact, the strand model adds a further, but related number 3 to this list, namely the maximum number of strands that make up elementary particles.

Page 209 The three-dimensionality of space is, as we saw already above, a result of the existence of strand tangles: only three dimensions allow tangles of strands. In short, all numbers 3 that appear in fundamental physics are explained by strands.

Predictions about dark matter and searches for new physics

Following the vast majority of scholars, astrophysical observations imply that galaxies and galaxy clusters are surrounded by large amounts of matter that does not radiate. This unknown type of matter is called *dark matter*.

In the strand model, the known elementary particles are the only possible ones. Therefore,

▷ The strand model predicts that dark matter, if it exists, is a mixture of particles of the standard model and black holes.

Page 164 This statement settles a further item from the millennium list of open issues.

The prediction from 2008 of a lack of new elementary particles in dark matter is at odds with the most favoured present measurement interpretations, but cannot yet be ruled out. The detection of black hole mergers in 2015 can even be seen as a partial confirmation. However, the issue is obviously not yet settled. In fact, the prediction provides another hard test of the model: if dark matter is found to be made of yet unknown particles, the strand model is in trouble.

We can condense all the results on particle physics found so far in the following statement:

▷ There is nothing to be discovered about nature outside general relativity and the standard model of particle physics.

Strands predict that there is no hidden aspect of nature left. In particular, the strand model predicts a so-called high-energy *desert*: it predicts the lack of any additional elementary particle. Equivalently, the strand model predicts that apart from the Planck scale, there is no further energy scale in particle physics. Some researchers call this beautiful result the *nightmare scenario*.

In other words, there is no room for discoveries beyond the Higgs boson at the Large Hadron Collider in Geneva, nor at the various dark matter searches across the world. If any new elementary particle is discovered, the strand model is wrong. More precisely, if any new elementary particle *that contradicts the strand model* is discovered, the strand model is wrong. That some unknown elementary particle has been missed in the present exploration of tangle classes is still a logical possibility, however.

Because the strand model confirms the standard model and general relativity, a further prediction can be made: *the vacuum is unique and stable.* There is no room for other options. For example, there are no domains walls between different vacuum states and the universe will not decay or change in any drastic manner.

In summary, the strand model predicts a lack of any kind of science fiction in modern physics.

Chapter 12
PARTICLE PROPERTIES DEDUCED FROM STRANDS

> Tutto quel che vedete, lo devo agli spaghetti.**
>
> Sophia Loren

The Planck units, via strands and the fundamental principle, explain almost all that is known about motion: strands explain *what* moves and *how* it moves. But the strand model is only correct if it also explains every measured property of every elementary particle. So far, we only deduced the spectrum and the quantum numbers of the elementary particles. Three kinds of particle properties from the millennium list remain open: the *masses*, the *mixing angles* and the *couplings*. These measured particle properties are important, because they determine the amount of change – or physical action – induced by the motion of each elementary particle.

So far, the strand model has answered all open questions on motion that we explored. In particular, the strand model has explained why quantum field theory, the interactions, the particle spectrum, general relativity and cosmology are what they are. But as long as we do not understand the measured properties of the elementary particles, we do not understand motion completely.

In short, the next step is to find a way to *calculate* these particle properties – and obviously, to show that the calculations agree with the measurements. The step is particularly interesting; so far, no other unified model in the research literature has ever achieved such calculations – not even calculations that disagree with measurements.

Because the strand model makes no experimental predictions that go beyond general relativity and the standard model of particle physics, explaining the properties of elementary particles is the *only way* to confirm the strand model. Many ways to test or to refute the strand model are possible; but only a calculation of the measured particle properties can confirm it.

The ideas in this chapter are more speculative than those of the past chapters, because the reasoning depends on the way that specific tangles are assigned to specific particles. Such assignments are never completely certain. We continue keeping this in mind.

** 'Everything you see, I owe it to spaghetti.' Sofia Villani Scicolone is an Italian actress and Hollywood star.

TABLE 14 The measured elementary particle masses, as given by the Particle Data Group in 2016.

Elementary particle	Mass value
Electron neutrino	$< 2\,\mathrm{eV}/c^2$
Muon neutrino quark	$< 2\,\mathrm{eV}/c^2$
Tau neutrino	$< 2\,\mathrm{eV}/c^2$
Electron	$0.510\,998\,9461(31)\,\mathrm{MeV}/c^2$
Muon	$105.658\,3745(24)\,\mathrm{MeV}/c^2$
Tau	$1776.86(12)\,\mathrm{MeV}/c^2$
Up quark ($q = 2/3\,e$)	$2.2(6)\,\mathrm{MeV}/c^2$
Down quark ($q = -1/3\,e$)	$4.7(5)\,\mathrm{MeV}/c^2$
Strange quark ($q = -1/3\,e$)	$96(8)\,\mathrm{MeV}/c^2$
Charm quark ($q = 2/3\,e$)	$1.27(3)\,\mathrm{GeV}/c^2$
Bottom quark ($q = -1/3\,e$)	$4.18(4)\,\mathrm{GeV}/c^2$
Top quark ($q = 2/3\,e$)	$173.21(1.22)\,\mathrm{GeV}/c^2$
W boson	$80.385(15)\,\mathrm{GeV}/c^2$
Z boson	$91.1876(21)\,\mathrm{GeV}/c^2$
Higgs boson	$125.09(24)\,\mathrm{GeV}/c^2$
Photon	not detectable
Gluons	not detectable
Graviton	not detectable
For comparison: the corrected Planck mass $\sqrt{\hbar c/4G}$	$0.611 \cdot 10^{19}\,\mathrm{GeV}/c^2$

THE MASSES OF THE ELEMENTARY PARTICLES

The mass describes the inertial and gravitational effects of a body. The strand model must reproduce all mass values observed in nature; if it doesn't, it is wrong.

To reproduce the masses of *all* bodies, it is sufficient that the strand model reproduces the measured masses, the mixing angles and the coupling strengths of the *elementary* particles. We start with their masses. All measured mass values are given in Table 14. All these values – more precisely: their ratios to the Planck mass – are unexplained and are part of the millennium list of open issues in fundamental physics.

In nature, the *gravitational mass* of a particle is determined by the space curvature that it induces around it. In the strand model, this curvature is due to the modified fluctuations that result from the presence of the tangle core; in particular, the curvature is due to the modified fluctuations of the particle tails – twisted tail pairs – and to the modified vacuum strand fluctuations just around the particle position. The modified strand shape fluctuations produce a crossing switch distribution around the tangle core; the crossing switch distribution leads to spatial curvature; at sufficiently large distances, this curvature distribution is detected as a gravitational mass.

In contrast, *inertial mass* appears in the Dirac equation. In the strand model, inertial mass is determined by the frequency and the wavelength of the helix drawn by the rotating phase vector. These quantities in turn are influenced by the type of tangle, by the fluctuations induced by the particle charges, by the topology changes induced by the weak interaction, and, in the case of fermions, by the average frequency and size of the belt (and possibly leather) trick. All these processes are due to strand shape fluctuations.

In short, both gravitational and inertial particle mass are due to strand fluctuations. More specifically, the mass seems mainly due to the fluctuations of the *tails* of the particle tangle: gravitational and inertial mass are due to the belt trick. *The strand model thus suggests that gravitational and inertial mass are automatically equal.* In particular, the strand model suggests that every mass is surrounded by fluctuating tails with crossing switches whose density decreases with distance and is proportional to the mass itself. As discussed above, this idea leads to universal gravity. Page 282

General properties of particle mass values

So far, our adventure allows us to deduce several results on the mass values of elementary particles:

— The strand model implies that the masses of elementary particles are *not free parameters*, but that they are determined by the specific topology, or tangledness, of the underlying tangles and their tangle families. Particle masses are thus *fixed* and *discrete* in the strand model – as is observed. Of course, we have to take into account all the members in each tangle family.
— The strand model implies that masses are always *positive* numbers.
— The strand model implies that the *more complex* a tangle is, the *higher* its mass value is. This follows from the behaviour of tangle tail fluctuations around the tangle core.
— Because particle masses are due to strand fluctuations, the strand model also implies that all elementary particle masses are *much smaller than the Planck mass*, as is observed. Also this result follows from the behaviour of tangle tail fluctuations around the tangle core.
— Because particle masses are due to strand fluctuations, particle and antiparticle masses – their tangles are mirrors of each other – are always *equal*, as is observed.
— Because particle masses are due to strand fluctuations, particle masses do *not* depend on the age of the universe, nor on their position in the universe, nor on any other state variable: The strand model predicts that particle masses are constant and invariant, as is observed. Page 373
— Because particle masses are due to strand fluctuations, and the fluctuations differ somewhat for tight and loose tangles of the same shape and topology, the strand model predicts that particle masses change – or *run* – with energy, as is observed.

The general properties of particle masses are thus reproduced by the strand model. Therefore, continuing our exploration makes sense. We start by looking for ways to determine the mass values from the tangle structures. We discuss each particle class separately, first looking at mass ratios, then at absolute mass values.

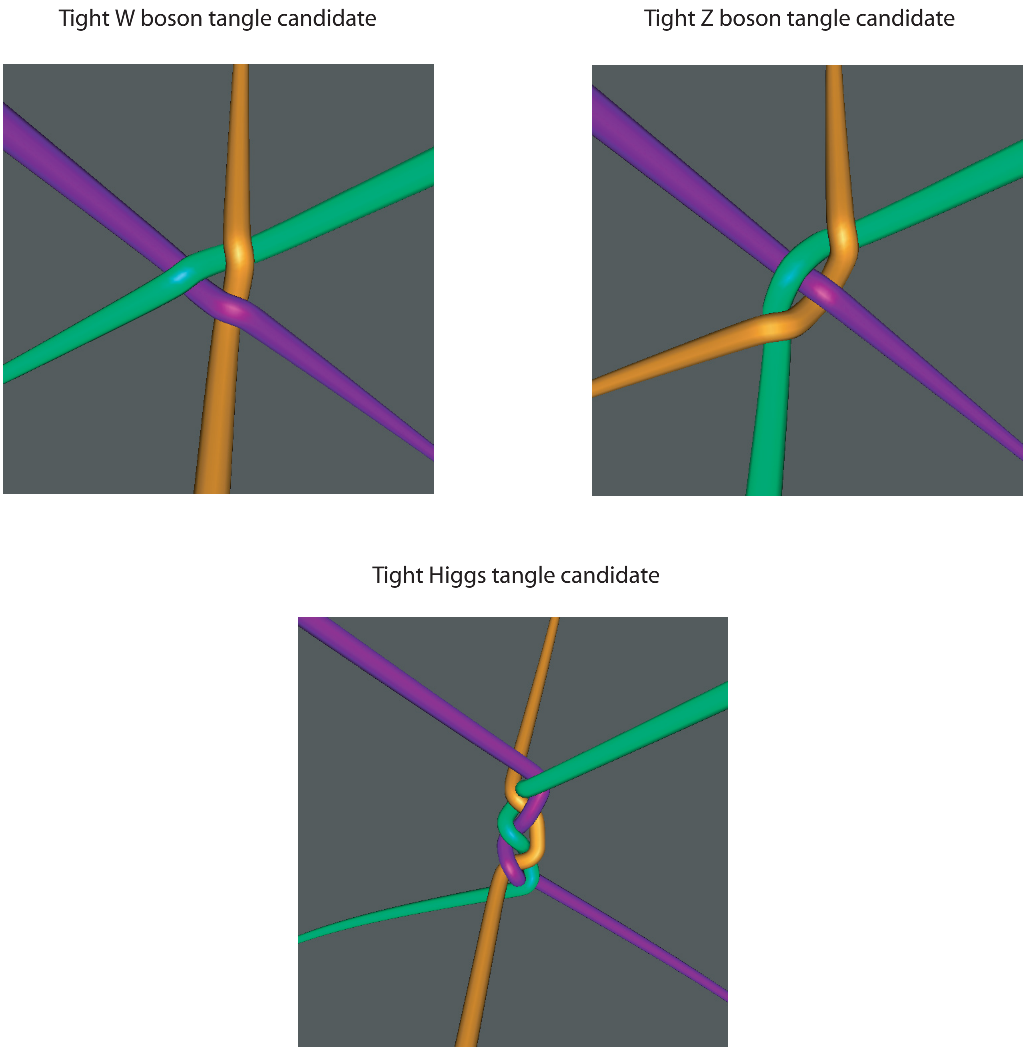


FIGURE 112 Tight tangle candidates (of 2015/2016) for the simplest tangles of the W, the Z and the Higgs bosons. In contrast to the pictures, the W and Z tails lie in a plane.

Boson masses

Three elementary particles of integer spin have non-vanishing mass: the W boson, the Z boson and the Higgs boson. Mass calculations are especially simple for bosons, because in the strand model, they are *clean* systems: each boson is described by a relatively simple tangle family; furthermore, bosons do not need the belt trick to rotate continuously.

We expect that the induced curvature, and thus the gravitational mass, of an elementary boson is due to the disturbance it introduces into the vacuum. *At Planck energy*, this disturbance will be, to a large extent, a function of the *ropelength* introduced by the corresponding *tight* tangle. Let us clarify these concepts.

Tight or *ideal* tangles or knots are those tangles or knots that appear if we imagine strands as being made of a rope of *constant* diameter that is *infinitely flexible*, *infinitely* *slippery* and pulled as tight as possible. Examples of tight tangles are shown in Figure 112. With physical ropes from everyday life, tight knots and tangles can only be approximated, because they are not infinitely flexible and slippery; tight tangles are mathematical idealizations. But tight tangles of strands are of special interest: if we recall that each strand has an effective diameter of one Planck length, tight tangles realize the Planck limit of the strand model.

— The *ropelength* of a tight *closed knot* is the length of a perfectly flexible and slippery rope of constant diameter required to tie the tight knot. In other words, the ropelength is the smallest amount of idealized rope needed to tie a knot.
— The *ropelength* of a tight *open knot* is the length by which a very long rope tied into a tight knot is shortened.
— With a bit of care, the concept of ropelength can be also be defined for tangles of several strands.

In the following, the ropelength is assumed to be measured in units of the rope *diameter*. Measuring ropelength in units of the rope radius is less common.

In the strand model, the ropelength measures, to a large extent, the amount by which a tight knot or tangle disturbs the vacuum around it. The ropelength fulfils all the properties of particle mass mentioned above: the ropelength is discrete, positive, increases with tangle complexity, is equal for particles and antiparticles, and is a constant and invariant quantity. The ropelength will thus play an important role in any estimate of a particle mass.

Page 359

It is known from quantum field theory that the masses of W and Z bosons do not change much between Planck energy and everyday energy, whatever renormalization scheme is used. This allows us, with a good approximation, to approximate the weak boson masses at low, everyday energy with their mass values at Planck energy. Thus we can use tight tangles to estimate boson masses.

Ref. 244

In the strand model, the *gravitational mass* of a spin 1 boson is proportional to the radius of the disturbance that it induces in the vacuum. For a boson, this radius, and thus the mass, scales as the third root of the ropelength of the corresponding tight tangle.

W/Z boson mass ratio and mixing angle (in the 2016 tangle model)

Page 360

Candidates for the simplest tangles of the W boson and of the Z boson families are shown in Figure 112. The corresponding ropelength values for tight tangles, determined numerically, are $L_W = 4.28$ and $L_Z = 7.25$ rope diameters. The strand model estimates the W/Z mass ratio by the cube root of the ropelength ratio:

Ref. 246

$$\frac{m_W}{m_Z} \approx \left(\frac{L_W}{L_Z}\right)^{1/3} = 0.84\ . \tag{196}$$

Ref. 233

This value has to be compared with the experimental ratio of 80.4 GeV/91.2 GeV=0.88. The agreement between experiment and strand model is not good. But the result is acceptable. First of all, the strand model reproduces the higher value of the neutral Z boson's mass: a tangle with spatial symmetry is more complex than one without. Finally, it is also clear why the calculated mass ratio does not match the experimental result.

First, the simple tangles represent and approximate W and Z bosons only to the first order. Page 253 As mentioned above, in the strand model, every massive particle is represented by an infinite family of tangles. The strand model thus also predicts that the match between the calculated and the measured ratio m_W/m_Z should improve when higher-order Feynman diagrams, and thus more complicated tangle topologies, are taken into account. Improving the calculation is still a subject of research. Secondly, approximating the tight knot effects with an effective radius, thus just using the ropelength to determine the mass, implies neglecting the actual shape, and effectively approximating their shape by a sphere. Thirdly, as already mentioned, this calculation assumes that the low energy mass ratio and the mass ratio at Planck energy are equal.

Despite the used approximations, the tight tangle estimate for the W/Z mass ratio gives an acceptable agreement with experiment. The main reason is that we expect the strand fluctuations from the various family members to be similar for particles with the *same* number of strands. For these mass ratios, the tail braiding processes cancel out. Also the other two approximations are expected to be roughly similar for the two weak bosons. This similarity explains why determining the W/Z boson mass *ratio* is possible with acceptable accuracy.

The W/Z mass ratio also determines the weak mixing angle θ_w of the weak interaction Lagrangian, through the relation $\cos\theta_w = m_W/m_Z$. The strand model thus predicts the value of the weak mixing angle to the same accuracy as it predicts the W/Z mass ratio.

This argument leads to a puzzle: Can you deduce from the strand model how the W/Z mass ratio changes with energy? Challenge 204 ny

Also the *inertial masses* of the W and Z bosons can be compared. In quantum theory, the inertial mass relates the wavelength and the frequency of the wave function. In the strand model, a quantum particle that moves through vacuum is a tangle core that rotates while advancing. Page 196, page 215 The frequency and the wavelength of the helix thus generated determine the inertial mass. The process is analogous to the motion of a body moving at constant speed in a viscous fluid at small Reynolds numbers. Despite the appearance of friction, the analogy is possible. If a small body of general shape is pulled through a viscous fluid by a constant force, such as gravity, it follows a *helical* path. Ref. 163 This analogy implies that, for spin 1 particles, the frequency and the wavelength are above all determined by the effective radius of the small body. This radius is the main influence on the frequency of the belt trick. The strand model thus suggests that the inertial mass – inversely proportional to the path frequency and the path wavelength squared – of the W or the Z boson is *approximately* proportional to its tight knot radius. This radius is given by the cube root of the ropelength. We thus get the same result as for the gravitational mass.

Also the inertial mass is not exactly proportional to the average tight knot radius; the precise shape of the tight knot and the other tangle family members play a role. The strand model thus predicts that a more accurate mass calculation has to take into account these effects.

In summary, the strand model predicts, from the belt trick, a W/Z mass ratio and thus a weak mixing angle close to the observed ratio, and explains the deviation of the approximation from the measured value – provided that the tangle assignments are correct.

The g-factor of the W boson

Ref. 233 Experiments show that the W boson has a g-factor with the value $g_W = 2.2(2)$. The limited accuracy does not yet allow to detect any anomalous magnetic moment, which would be especially interesting. Nevertheless, the results can be compared to the prediction of the strand model.

The strand model makes a simple prediction for charged elementary particles: because mass rotation and charge rotation are both due to the rotation of the tangle core, the g-factor of all such particles is 2 – in the approximation that neglects Feynman diagrams of higher order, i.e., that neglect anomalous effects. In particular, the g-factor of the W boson is predicted to be 2 in this approximation. Also this prediction thus agrees both with experiment and with the standard model of particle physics. Ref. 245

The Higgs/Z boson mass ratio

The observed mass value of the Higgs boson is 125(1) GeV. The observed mass value for the Z boson is 91.2(1) GeV. Like for the other bosons, the strand model suggests using the ropelength to estimate the mass of the Higgs boson tangle. The candidate tangle for the Higgs boson was illustrated above in Figure 96, and its tight version is shown in Figure 112. Page 331 Page 360

Ref. 246 The ropelength of the tight Higgs tangle turns out to be 17.1 diameters, determined by Eric Rawdon with a computer approximation. This value yields a naive mass estimate for the Higgs boson of $(17.1/7.25)^{1/3} \cdot 91.2$ GeV, i.e.,

$$m_{\text{Higgs}} \approx 121 \text{ GeV} . \qquad (197)$$

Starting with the W boson yields an estimate for the Higgs mass of 128 GeV. Both estimates are not good but acceptable, given that the non-sphericity of the W, Z and Higgs boson tangles have not been taken into account. (The strand model suggests that for a strongly non-spherical shape – such as the shape of the W, Z and Higgs tangle – the effective mass is higher than the value deduced from ropelength alone.) Deducing better mass ratio estimates for the W, Z and Higgs tangles is still a subject of research.

In summary, the strand model predicts Higgs/Z, Higgs/W and W/Z mass *ratios* close to the observed values; and the model suggests explanations for the deviations of the approximation from the observed value – provided that the tangle assignments for the three bosons are correct.

A first approximation for absolute boson mass values

The tangles for the W, Z and Higgs bosons also provide a first approximation for their *absolute* mass values. The tangles are rational; in particular, each tangle is made of strands that can be pulled straight. This implies, for each strand separately, that a configuration with no extra strand length and no net core rotation is possible. As a result, in the first approximation, the gravitational mass and the inertial mass of the elementary bosons both *vanish*.

A better approximation for mass values requires to determine, for each boson, the probability of crossing switches in and around its tangle core. This probability depends

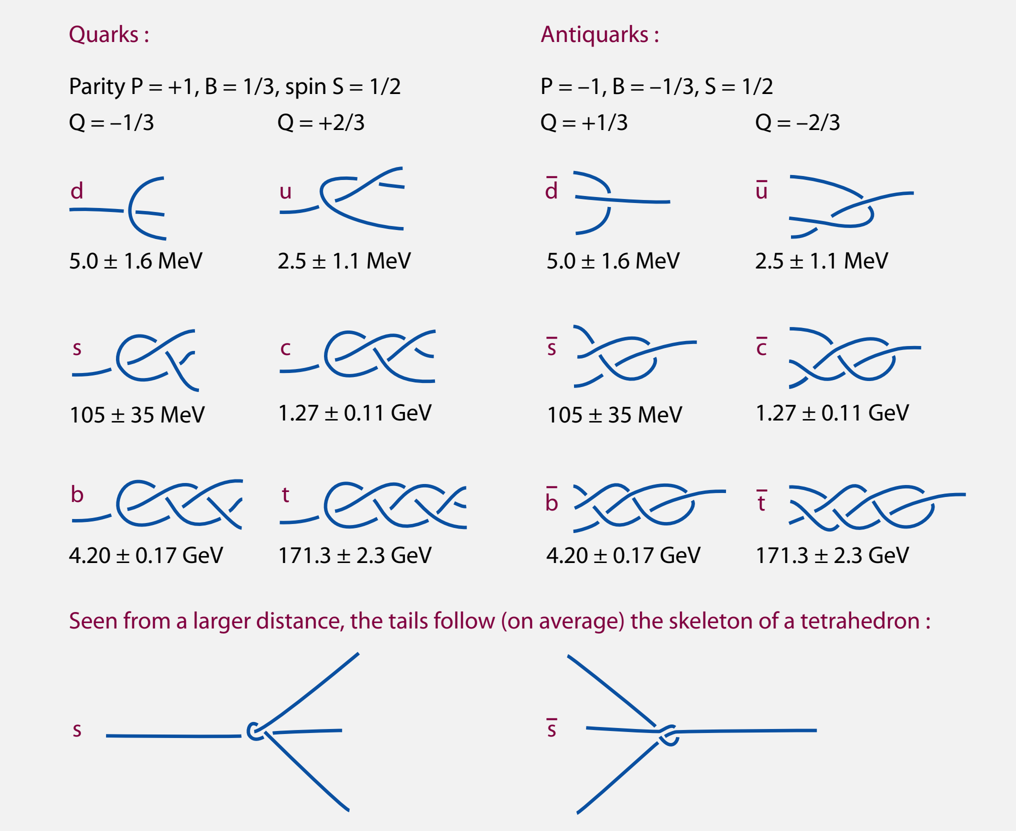


FIGURE 113 The simplest tangles assigned to the quarks and antiquarks. The experimental mass values are also given. Calculated mass values are still a topic of research.

on the probabilities for tail braiding and for core rotation. These probabilities are low, because, sloppily speaking, the corresponding strand fluctuations are rare. The rarity is a due to the specific tangle type: tangles whose strands can be pulled straight have low crossing switch probabilities at their core or at their tails when they propagate.

The strand model thus predicts that elementary boson masses, like all other elementary particle masses, are much smaller than the Planck mass, though not exactly zero. This prediction agrees with observation: experimentally, the mass values for the W, Z and Higgs are of the order of 10^{-17} Planck masses. We will search for more precise mass estimates below.

Quark mass ratios

Quarks are fermions. In the strand model, mass estimates for fermions are more difficult than for bosons, because their shapes and their tails are less symmetric. Still, using Figure 113, the strand model allows two predictions about the relations between quark masses.

— The quark masses are predicted to be the same for every possible colour charge. This is observed.

Page 321
— The progression in ropelength of the tight basic tangles for the six quarks suggests a progression in their masses. This is observed, though with the exception of the up quark mass. For this exceptional case, effects due to tail braiding, to the Higgs boson, and to quark mixing are expected to play a role, as argued below. Page 376

Let us try to extract numerical values for the quark mass ratios. We start by exploring the tight quark tangles, thus Planck-scale mass values. For each quark number q, the quark mass will be the weighted average over the mass of its family tangles with q, $q+6$, $q+12$, ... crossings, where the period 6 is due to the Higgs boson. Each tight tangle has a certain ropelength. The mass of each tangle will be determined by the frequency of crossing changes at the core, including those due to the belt trick. The quark mass then is the average over all family tangles; it will be determined by the frequency of tail braiding and of all other fluctuations that generate crossing switches.

For determining mass ratios, the frequency of the crossing switches at the core are the most important. Given that the particles are fermions, not bosons, this frequency is expected to be an exponential of the ropelength L. Among quarks, we thus expect a general mass dependence of the type

$$m \sim \mathrm{e}^{aL} \qquad (198)$$

where a is an unknown number of order 1. We note directly that such a relation promises general agreement with the observed quark mass ratios. Nevertheless, the exponential dependence must not be seen as more than an educated guess.

Ref. 247
Actual ropelength calculations by Eric Rawdon and Maria Fisher show that the ropelength of quark tangles increases roughly linearly with q, as expected from general knot theoretic arguments. Their results are given in Table 15. Comparing these calculated ropelength differences with the known low-energy quark masses confirms that the number a has an effective value in the range between 0.4 and 0.9, and thus indeed is of order one.

The results of Table 15 suggest that the top quark should be particularly heavy – as is observed. The results of Table 15 also suggest that something special is going on for the u-d quark pair, which is out of sequence with the other quarks. Indeed, the strand model predicts a very small mass, – at the Planck scale – for the down quark. However, in nature, the down mass is observed to be *larger* than the up mass. (We note that despite this issue, meson mass sequences are predicted correctly.)

It could well be that the large symmetry of the simplest down quark tangle is the reason for the exceptional mass ordering. The large symmetry of the down tangle should lead to easier, i.e., more frequent interactions with the Higgs boson. In other words, the braiding, i.e., the mixing with the more massive family members, appears to be *higher* for the down quark than that for the up quark, and this would explain the higher mass of the down quark. Maybe this explanation can be checked with data.

The experimental values for the quark masses are given in Table 16; the table also includes the values extrapolated to Planck energy for the pure standard model. The calculation of the strand model *does not agree* with the data. The only encouraging aspect is that the ropelength approximation provides an approximation for older speculations Ref. 248

TABLE 15 Calculated ropelengths, in units of the rope *diameter*, of tight quark tangles of Figure 86 (Page 320) with tails oriented along the skeleton of a tetrahedron.

Tangle	Length	Ropelength	Difference
skeleton (vacuum)	138.564065	base value	
simplest d	139.919533	1.355468	1.355468
simplest u	142.627837	4.063773	2.708305
simplest s	146.175507	7.611443	3.547670
simplest c	149.695643	11.131578	3.520136
simplest b	153.250364	14.686299	3.554721
simplest t	157.163826	18.599761	3.913462

TABLE 16 For comparison: the quark masses at Planck energy,, calculated from the measured quark masses using the standard model of particle physics – assuming that it is correct up to Planck energy.

Quark	Low energy mass	Planck energy mass
u ($q = 2/3e$)	2.5(1.1) MeV	0.45(0.16) MeV
d ($q = -1/3e$)	5.0(1.6) MeV	0.97(0.10) MeV
s ($q = -1/3e$)	105(35) MeV	19.4(1.2) MeV
c ($q = 2/3e$)	1270(110) MeV	213(8) MeV
b ($q = -1/3e$)	4200(170) MeV	883(10) MeV
t ($q = 2/3e$)	171300(2300) MeV	66993(880) MeV

on approximately *fixed mass ratios* between the up-type quarks u, c, t and *fixed mass ratios* between down-type quarks d, s, b. The attempted strand model estimate shows that ropelength alone is *not sufficient* to understand quark mass ratios. Research has yet to determine which tangle shape aspect has to be included to improve the correspondence with experiment.
Ref. 249

In fact, the strand model predicts that everyday quark masses result from a combination of three effects: the effect of ropelength and of tangle core shape on rotation and the belt trick, the effect of sixfold tail braiding, and the effect of the energy dependence of mass between Planck energy and everyday energy, due to core loosening.

In short, an analytic calculation for quark masses seems difficult, due to their elongated shape. Better analytical approximations should be possible, though. And with sufficient computer power, it will also be possible to determine the frequency of core shape deformations and core rotations, including the belt trick. It will also be possible to determine the energy dependence of the quark masses, and the probability for tail braiding. More research is needed on all these points; in the end, it will allow to calculate quark masses.

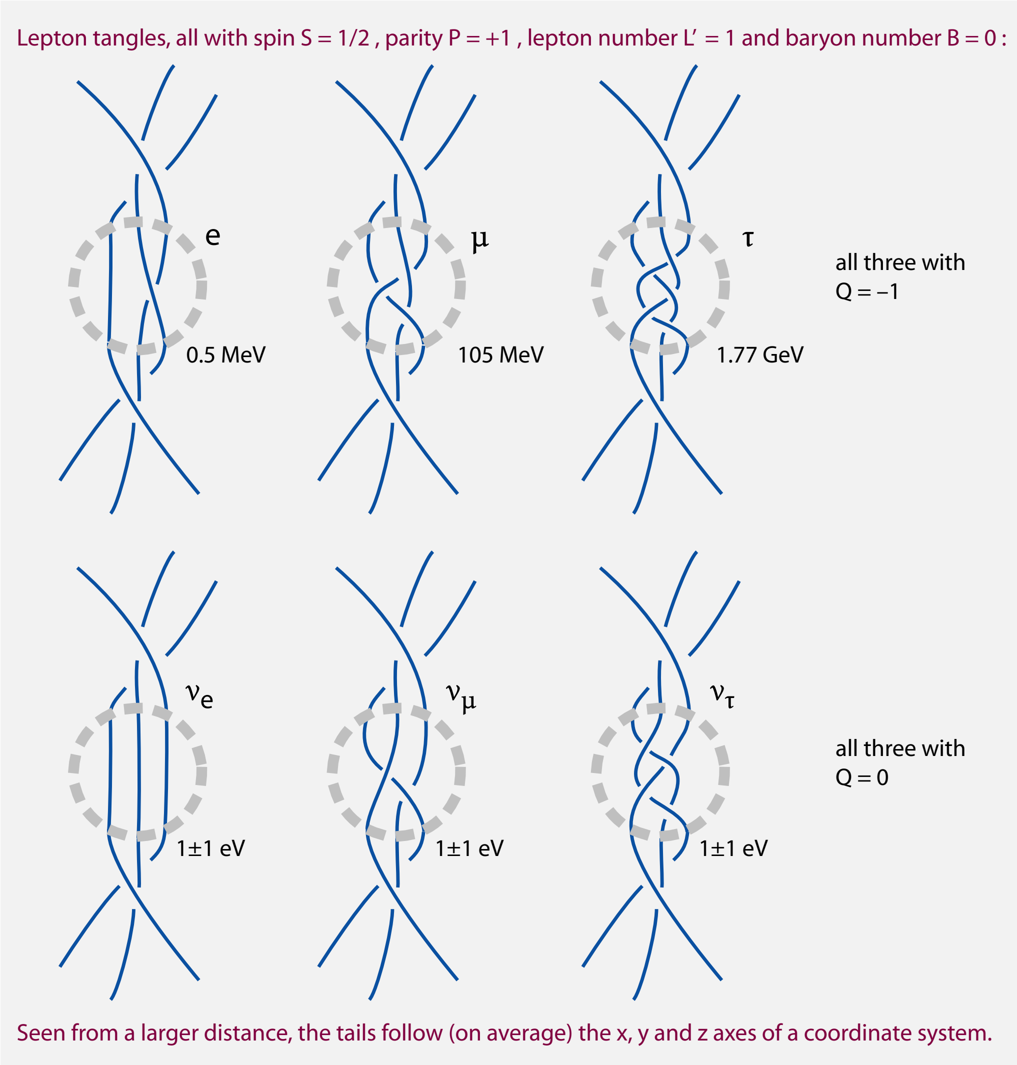


FIGURE 114 Simple – wrong – candidate tangles for the leptons. Antileptons are mirror tangles. The experimental mass values are also given. These wrong tangles yield the wrong mass sequence.

Lepton mass ratios

Mass calculations for leptons are as involved as for quarks. Each lepton, being a fermion, has a large family of associated tangles: there is a simplest tangle and there are the tangles that appear through repeated application of tail braiding. Despite this large tangle family, some results can be deduced from the simplest lepton tangles alone, disregarding the higher-order family members.

Mass calculations require first to find the correct lepton tangles. This is not straightforward. A consistent proposal appears when the set of lepton tangles reproduces all Feynman diagrams. This proposal is shown in Figure 115.

Both for neutrinos and for charged leptons, the progression in ropelength of the tight

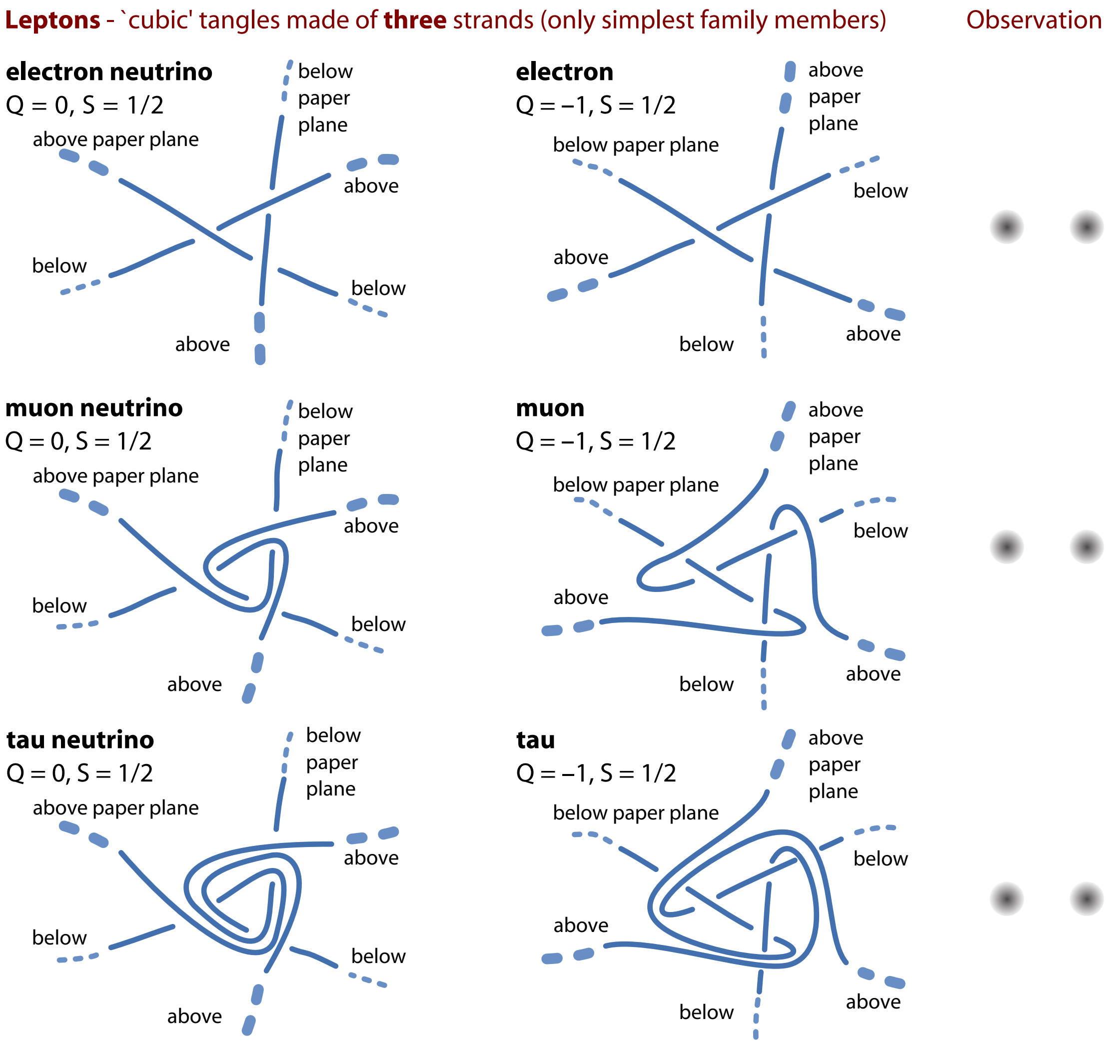


FIGURE 115 Presently proposed tangles for the leptons. Antileptons are mirror tangles. Feynman diagrams are reproduced correctly. The handedness of neutrinos is apparent. Mass sequences are as observed. The experimental mass values are also listed. At present, estimated mass values are within a few orders of magnitude.

versions of the basic tangles predicts a progression in their masses. This is indeed observed. In addition, the tangles of Figure 115 imply that the neutrinos are much less massive than the charged leptons. Spin 1/2 arises. Also the electric charge quantum number arises, as long as charge is related to topological chirality. And the handedness of neutrinos is evident. These tangle assignments are thus promising.

For each lepton tangle with l crossings, knot theory predicts a ropelength L that increases roughly proportionally to the crossing number: $L \sim l$. Each lepton mass value will again be given by the frequency of crossing switches due to rotations, including the belt trick, and of tail braiding. We thus expect a general relation of the type

$$m_l \sim e^{bL_l} \tag{199}$$

where b is a number of order 1 that takes into account the shape of the tangle core. Such a relation is in general agreement with the observed ratios between lepton masses.

We note that the lepton mass generation mechanism of the strand model differs from other proposals in the research literature. It agrees with the Higgs mechanism but goes beyond it. For neutrinos, the mechanism contradicts the see-saw mechanism but confirms the Yukawa mechanism directly. From a distance, the mass mechanism of the strand model also somewhat resembles conformal symmetry breaking.

Page 330
Ref. 250

In sort, research on lepton masses is ongoing; calculations of ropelengths and other geometric properties of the lepton tangles will allow a more detailed analysis. The main challenge remains to estimate the neutrino masses before experiments – such as KATRIN – determine them.

On the absolute values of particle masses

In nature, the masses of elementary particles are observed to be much lower than the Planck mass: the observed values lie between about 10^{-30} for neutrinos and 10^{-17} for the top quark. Particle masses are constant over space and time. Antiparticles have the same mass as particles. Gravitational and inertial masses are the same. Following the standard model, particle masses are due to the Higgs mechanism. Finally, elementary particles masses run with energy.

All qualitative observations about particle mass are reproduced by the strand model. However, the explanation of the numerical values is still lacking.

In the strand model, the gravitational mass of elementary particles is due to disturbance of the vacuum, in particular to the disturbance of the vacuum fluctuations. Larger masses are due to more complex tangles. Since rest mass is localized energy, rest mass is due to crossing switches per time. Larger masses have more crossing switches per time than lower masses.

In the strand model, the inertial mass of elementary particles is their reluctance to rotate. Inertial mass describes the relation between rotation frequency and wavelength; in other terms, inertial mass described the steepness of the helix drawn by the rotating phase arrow of a propagating particle. Larger masses have low steepness, smaller masses have higher steepness. Larger masses are due to more complex tangles.

As we just saw, the strand model predicts mass *sequences* and *ratios* of elementary particle masses that corroborate or at least do not contradict observations too much. The next step is to determine *absolute* mass values from the strand model. So far we only found that elementary particle masses are much smaller than a Planck mass. But to validate the strand model, we need more precise statements.

To determine gravitational mass values, we need to count those crossing switches that occur at rest; to determine inertial mass values, we need to look for crossing switches in the case of a moving particle – or, if we prefer, to understand the origin of the steepness of the helix drawn by the phase arrow. All these methods should first lead to mass value estimates and then to mass value calculations.

In general, the strand model reduces mass determination to the calculation of the details of a process: How often do the fluctuations of strands lead to crossing switches? There are various candidates for the crossing switches that lead to particle mass.

The first candidate for mass-producing crossings is tail switching. In general however, tail switching leads to different particle types. Only for the Higgs process, i.e., the addition of a full Higgs braid to a particle, is this process expected to be relevant. We can also say that tail braid addition are the strand model's version of the Yukawa coupling terms.

The next candidate for mass-producing crossings is the belt trick. The belt trick emits twisted tether pairs, thus virtual gravitons, and is a good candidate to understand gravitational mass. The belt trick also leads to core rotation and core displacement, which is the essence of inertial mass. Determining probabilities for the belt trick thus promises to allow mass calculations.

The third candidate for mass-producing crossings could appear when particles shed one strand and grab a new one. The influence of this process is not clear yet.

A fourth candidate for mass-producing crossings is the leather trick. However, the leather trick cannot be realized for strands that reach spatial infinity; therefore it is expected that it plays no role. In fact, the supposed effects of the leather tricks in an early phase of the strand model have now become effects due to the Higgs absorption and emission.

A fifth candidate for mass-producing crossings are those crossings that occur *above* or *around* the core, similar to the crossing that occur above the horizon of a black hole. This candidate group includes the belt trick, includes the Higgs mechanism, and thus is equivalent to the set of previous candidates.

It might be that some other mass-producing switching processes are being overlooked, but this seems unlikely. Therefore, in the following we explore the fifth candidate group in more detail, namely the crossings *around* a given tangle core.

Before looking for estimates, we note that in the past, various researchers have reached the conclusion that all elementary particle masses should be due to a common process or energy scale. Among theoretical physicists, the breaking of conformal symmetry has always been a candidate for such a process. Among experimental physicists, the Higgs mechanism – now confirmed by experiment – is the favourite explanation of all elementary particle masses. In the strand model, crossing switches around tangles are related to the Higgs boson. At the same time we can also argue that tangles break the conformal symmetry of vacuum. With a bit of distance, we can thus say that the strand model agrees with both research expectations.

Ref. 250, Ref. 251
Page 330

Let us continue with the quest for absolute mass estimates. In the strand model, *absolute* mass values are *not* purely geometric quantities that can be deduced directly from the shapes of tangle knots. Particle masses are due to *dynamical* processes. Absolute mass values are due to strand fluctuations; and these fluctuations are influenced by the core topology, the core shape, the core ropelength and core tightness.

To determine absolute particle mass values, we need to determine the ratio between the particle mass and the *Planck mass*. This means to determine the ratio between the crossing switch probability for a given particle and the crossing switch probability for a Planck mass, namely one switch per Planck time.*

* What is a Planck mass? In the strand model, a Planck mass corresponds to a structure that produces one

Energy is action per time. Mass is localized energy. In other words, the *absolute* mass of a particle is given by the average number of crossing switches it induces per time:

▷ Mass is crossing switch rate.

More precisely, the crossing switch rate of a particle at rest is its gravitational mass, and the crossing switch rate induced by propagation is its inertial mass. Let us explore the relations.

Given that mass is determined by the crossing switch rate, we deduce that particle mass values are determined by tangle topology, are fixed, are discrete, are positive, increase with tangle core complexity, are identical for particle and antiparticles, are constant over time, and are much smaller than the Planck mass. Because all these properties match observations, the local crossing switch rate indeed realizes all qualitative requirements for absolute particle mass values. We can thus proceed with the hope to learn more. In order to calculate absolute particle masses, we just need to determine the number of crossing switches per time that every particle tangle induces.

Page 359

One general way to perform a particle mass calculation is to use a *computer*, insert a strand model of the fluctuating vacuum plus the strand model of the particle under investigation, and count the number of crossing switches per time. The basis for one such approach, using the analogy of the evolution of a polymer in liquid solution, is shown in Figure 116. In contrast to polymers, also the change of strand length has to be taken into account. By determining, for a given core topology, the average frequency with which crossing switches appear for a tethered core, we can estimate the masses of the leptons, quarks and bosons. In such a mass calculation, the mass scale is set indirectly, through the time scale of the fluctuation spectrum. This is tricky but feasible. One would first need to find the parameter space and the fluctuation spectrum for which the polymer tangle follows the Schrödinger equation. Calculations with different tangles should then yield the different mass values. Such a simulation would also of interest for exploring the strand model of quantum mechanics.

A more precise computer simulation would also model the vacuum itself with strands. This approach would even allow to explore gravitational and inertial mass separately. In such a simulation, the particle mass appears when the helical motion of a tangle moving through a strand vacuum is observed. The required effort can be reduced by using the most appropriate computer libraries.

A further general way to determine particle masses is to search for *analytical approximations*. This is a fascinating conceptual and mathematical challenge. The main issue is to clarify which crossing switches contribute most to particle mass.

crossing switch for every Planck time, constantly, without interruption. But the strand model predicts that such structures do *not* appear as localized particles, because every localized particle – i.e., every tangle – has, by construction, a much smaller number of induced crossing switches per time. Following the strand model, elementary particles with Planck mass *do not exist*. This conclusion agrees with observation. But the strand model also implies that black holes with a Planck mass *do not exist*. Indeed, such Planck-scale black holes, apart from being extremely short-lived, have no simple strand structure. We can state that a Planck mass is never localized. Given these results, we cannot use a model of a *localized* Planck mass as a unit or a benchmark to determine particle masses.

The impossibility of using Planck mass as a unit is also encountered in everyday life: no mass measurement in any laboratory is performed by using this unit as a standard.

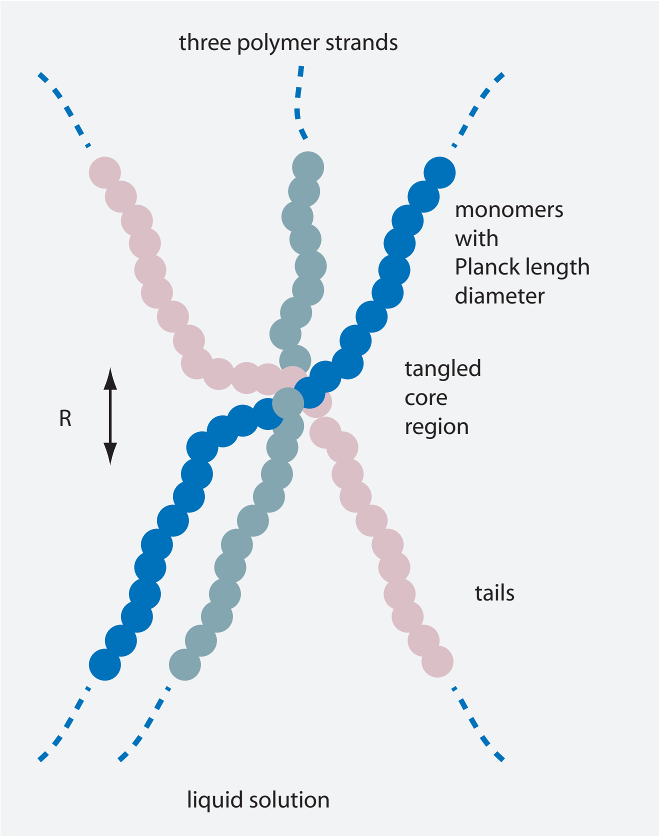


FIGURE 116 Determining lepton mass values with the help of a polymer analogy of strands. After rescaling, the probability of crossing switches around the tangle core yields an estimate for the mass of the elementary particle with that tangle.

Analytical estimates for particle masses

A first analytical attempt is the following. We assume that the inertial mass for a moving fermion is proportional to the fluctuation-induced appearance of the belt trick. If the tight core has a diameter of, say, three Planck lengths – and thus a circumference of around 9 Planck lengths – then the probability p of the belt trick for a particle with six tails will be in the range

$$p \approx (e^{-9})^6 \approx 10^{-24} . \qquad (200)$$

This value would be the order of magnitude for the mass estimate, in Planck units. Such an estimate is only very rough, and the exponent can be quite different. Nevertheless, we do get an explanation for the large difference between the Planck mass and the typical fermion mass. A more precise analytical approximation for the belt trick probability – not an impossible feat – will therefore solve the so-called *mass hierarchy problem*. We thus want to know:

Challenge 205 r

▷ What is the numerical probability of the belt trick for a tethered core of given topology with fluctuating tails?

So far, several experts on polymer evolution have failed to provide even the crudest estimate for the probability of the belt trick in a polymer-tethered ball. Can you provide one?

A second analytical approach starts from the following question:

Challenge 206 r ▷ How often does a tail cross above the tangle core?

This question is loosely related to the previous one; in addition, this approach illustrates why complex cores have larger mass. The probability of such crossings, when squared, would be an estimate for the crossing switch rate, and thus for the particle mass. (There are additional details to the calculation.) We note directly that the number of tails will have a smaller impact on mass than the complexity of the tangle. So far, a reliable estimate for the crossing number, as a function of the tangle core properties, is still missing – even a crude one. Can you find one? May be the roughly linear relation observed between ropelength and (average) crossing number can be of help. Ref. 252

Open issues about mass calculations

Calculating absolute particle masses from tangle fluctuations, either numerically or with an analytical approximation, will allow the final check of the statements in this section. The strand model predicts that the resulting values will match experiments. For these calculations, it is essential that the tangle assignment for each elementary particle is correct. Finally, in 2019, the tangles for the W tangle and for the lepton tangles seem settled. Challenge 207 e

* *

Because the strand model predicts a lack of new physics beyond the standard model of particle physics, the calculation of neutrino masses, and thus their mass sequence, is one of the few possible *predictions* – in contrast to retrodictions – that are left over in the strand model.

* *

Challenge 208 s Is the mass of a tangle related to the vacuum density of strands?

* *

Challenge 209 s Do particle masses depend on the cosmological constant?

* *

The mass of an elementary particle does not depend on the spin direction. In particular, the W and Z bosons have equal longitudinal and transversal mass. The strand model does not allow an influence of spin orientation on mass.

* *

Challenge 210 s Can the concept of *total curvature* of a tangle help to calculate particle masses?

* *

Does the effect of tail braiding confirm the conjecture that every experiment is described by a small energy scale, determining the resolution or precision, and a large energy scale, less obvious, that determines the accuracy? Challenge 211 d

* *

If tail braiding is due to the weak interaction, and if the Higgs is a tail-braided vacuum, can we deduce that the Higgs interaction is a higher order effect of the weak interaction? Can we deduce a concrete experimental prediction from this relation?

On fine-tuning and naturalness

It has become fashionable, since about a decade, to state that the standard model of elementary particle physics is 'fine-tuned'. The term expresses several ideas. First of all, the extremely low value of the vacuum energy is not obvious when all the zero-point field contributions from the various elementary particles of the standard model are included. A low vacuum energy seems only possible if the masses and the particle types of the standard model are somehow interrelated. In other words, the term 'fine tuning' expresses, above all, the *lack of understanding* of the origin of the masses, mixings and coupling constants of elementary particles.

The term 'fine tuning' is also used to state that the universe would be *very different* if the fundamental constants would be different. But this statement lacks deep truth. In this usage, the term 'fine tuning' states that particle masses are *not* parameters that can be varied at will. In common usage, 'parameters' are variable constants; but the low value of the vacuum energy – as well as many other observations – shows that the masses of elementary particles *cannot* be varied without destroying the validity of the standard model of particle physics.

Some people suggest that 'fine-tuning' implies that the standard model of particle physics is 'unnatural', whatever this might mean in detail. Some even suggest that the parameters of the standard model lack any explanation. The strand model – but also common sense – show that this suggestion is false.

The strand model *naturally* has a low vacuum energy, because the unknotted strands of flat space naturally have a zero energy density, and the particle masses, mixings and coupling constants are not variable or random, but *naturally* unique and fixed in value. Any *correct* description of nature must be 'fine-tuned'. If the standard model would not be 'fine-tuned', it would not describe nature.

In short, the fashionable term 'fine-tuned' is equivalent to the terms 'unmodifiable' and 'hard to vary' that were discussed above. Page 166 All these terms highlight the lack of alternatives to the world as we observe it, the existence of explanations for the processes around us, and our ability to discover and grasp them. This is part of the wonders of nature. And the strand model makes those wonders apparent at the Planck scale.

Summary on elementary particle masses and millennium issues

The strand model implies that masses are dynamic quantities fixed by processes due to the geometric and topological properties of specific tangle families. As a result, strands explain why the masses of elementary particles are not free parameters, but fixed and unique constants, and why they are much smaller than the Planck mass by many orders of magnitude. Strands also reproduce all known qualitative properties of particle masses. Page 371

Strands provide estimates for a number of elementary particle *mass ratios*, such as m_W/m_Z and m_{Higgs}/m_W. Most quark and lepton mass sequences and first rough estimates of mass ratios appear to agree with experimental data. All hadron mass sequences are predicted correctly.

The strand model also promises to calculate *absolute mass values*, including their change or 'running' with energy. Particle masses are in the range of 10^{-20} times the Planck mass, near the observed mass values. In the future, more precise calculations will allow either improving the match with observations or refuting the strand model.

The results are encouraging for two reasons. First of all, no other unified model that agrees with experiment explains the qualitative properties of mass and mass sequences. Secondly, no research on statistical rational tangles exists; an understanding of the parameters of nature might be lacking because results in this research field are lacking.

Page 164 In the millennium list of open issues we have thus caught a glimpse of how to settle the origin of particle masses – though we have not calculated them yet. Because further interesting challenges are awaiting us, we continue nevertheless. In the next leg, we investigate how elementary particle states mix.

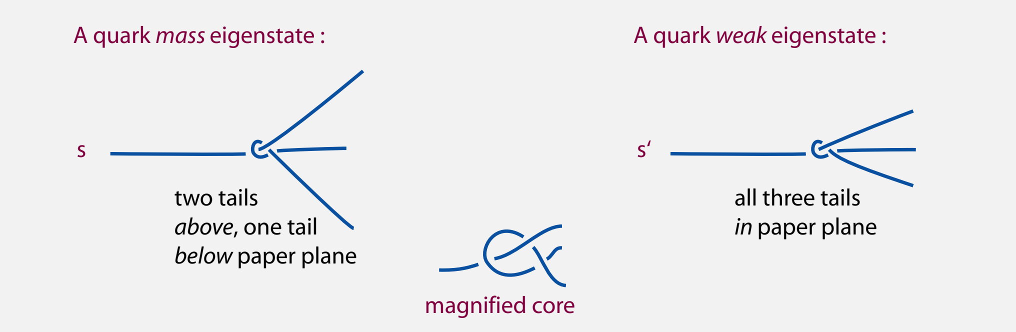


FIGURE 117 Tail shifting leads to quark mixing: mass eigenstates and weak eigenstates differ.

Mixing angles

In nature, the *mass* eigenstates for fermions differ from their *weak* eigenstates: quarks mix among themselves, and so do neutrinos. Quarks also show CP violation; for neutrinos, the issue is still open. These effects are described by two so-called *mixing matrices*. The two mixing matrices contain fundamental constants of nature. For the strand model to be correct, it must allow calculating the measured values of all components of the two mixing matrices.

Quark mixing – the experimental data

In nature, the quark mass eigenstates and their weak eigenstates differ. This difference was discovered in 1963 by Nicola Cabibbo and is called *quark mixing*. The values of the Ref. 233 elements of the quark mixing matrix have been measured in many experiments, and more experiments aiming to increase the measurement precision are under way.

Vol. V, page 251 The quark *mixing matrix*, also called CKM mixing matrix, is defined by

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = (V_{ij}) \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{201}$$

where, by convention, the states of the +2/3 quarks u, c and t are unmixed. Unprimed quarks names represent strong (and electromagnetic) eigenstates, primed quark names Ref. 233 represent weak eigenstates. In its standard parametrization, the mixing matrix reads

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{202}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$ and i and j label the generation $(1 \leqslant i, j \leqslant 3)$. The mixing matrix thus contains three mixing angles, θ_{12}, θ_{23} and θ_{13}, and one phase, δ. In the limit $\theta_{23} = \theta_{13} = 0$, i.e., when only *two* generations mix, the only remaining parameter is the angle θ_{12}, called the *Cabibbo angle*; this angle is Cabibbo's original discovery. The last parameter, the so-called *CP-violating phase* δ, by definition between 0 and 2π, is measured to be different from zero; it expresses the observation that CP invariance is violated in the case of the weak interactions. The CP-violating phase only appears in the third column of the matrix; therefore CP violation requires the existence of (at least) three generations.

Challenge 212 e

The present 90 % confidence values for the measured *magnitude* of the complex quark mixing matrix elements are

Ref. 233

$$|V| = \begin{pmatrix} 0.97427(14) & 0.22536(61) & 0.00355(15) \\ 0.22522(61) & 0.97343(15) & 0.0414(12) \\ 0.00886(33) & 0.0405(12) & 0.99914(5) \end{pmatrix} . \qquad (203)$$

All these numbers are unexplained constants of nature, like the particle masses. Within experimental errors, the matrix V is unitary.

A huge amount of experimental work lies behind this short summary. The data have been collected over many years, in numerous scattering and decay experiments, by thousands of researchers. Nevertheless, this short summary represents all the data that any unified description has to reproduce about quark mixing.

Quark mixing – explanations

In the standard model of particle physics, the quark mixing matrix is usually seen as due to the coupling between the vacuum expectation value of the Higgs field and the left-handed quark doublets or the right handed quark singlets. However, this description does not lead to a numerical prediction.

A slightly different description of quark mixing is given in the strand model. In the strand model, the Higgs field and its role as mass generator and unitarity maintainer is a special case of the process of tail braiding. And braiding is related to the weak interaction. Because the various quarks are differently tangled rational tangles, tail braiding can reduce or increase the crossings in a quark tangle, and thus change quark flavours. We thus deduce from the strand model that quark mixing is an automatic result of the strand model and related to the weak interaction. We also deduce that quark mixing is due to the *same* process that generates quark masses, as expected. But we can say more.

Page 330
Page 253

In the strand model, the *mass eigenstate* – and colour eigenstate – is the tangle shape in which colour symmetry is manifest and in which particle position is defined. The mass eigenstates of quarks correspond to tangles whose three colour-tails point in three directions that are equally distributed in space. The shape in which the tails point in three, equally spaced directions is the shape that makes the SU(3) representation under core slides manifest.

In contrast, the *weak eigenstates* are those shapes that makes the SU(2) behaviour of core pokes manifest. For a quark, the weak eigenstate appears to be that shape of a tangle for which all tails lie in a plane; for such plane configuration, the tails and the core mimic a belt and its buckle, the structure that generates SU(2) behaviour. The two types of eigenstates are illustrated in Figure 117.

In the strand model, masses are dynamical effects related to tangle shape. In the case of quarks, the two configurations just mentioned will thus behave differently. We call the transformation from a mass eigenstate to a weak eigenstate or back *tail shifting*. Tail shifting is a deformation: the tails as a whole are rotated and shifted. On the other hand, tail shifting can also lead to untangling of a quark tangle; in other words, tail shifting can lead to tail braiding and thus can transform quark flavours. The process of tail shifting can thus explain quark mixing. (Tail shifting also explains the existence of neutrino mixing, and the lack of mixing for the weak bosons.)

Tail shifting can thus be seen as a *partial* tail braiding; as such, it is due to the weak interaction. This connection yields the following predictions:

— Tail shifting, both with or without tail braiding at the border of space, is a generalized deformation. Therefore, it is described by a unitary operator. The first result from the strand model is thus that the quark mixing matrix is unitary. This is indeed observed. Page 210 Ref. 233
— For quarks, tail braiding is a process with small probability. As a consequence, the quark mixing matrix will have its highest elements on the diagonal. This is indeed observed.
— Tail shifting also naturally predicts that quark mixing will be higher between neighbouring generations, such as 1 and 2, than between distant generations, such as 1 and 3. This is also observed.
— The connection between mixing and mass also implies that the 1–2 mixing is stronger than the 2–3 mixing, as is observed.
— Finally, tail shifting predicts that the numerical values in the quark mixing matrix can be deduced from the difference between the shapes of the two kinds of tangles shown in Figure 117. In particular, tail shifting also predicts that the quark mixing angles change, or run, with energy. In addition, the effect is predicted to be small. On the other hand, so far there is no reliable experimental data on the effect.

Performing a precise calculation of mixing angles and their running with energy is still a subject of research.

A challenge

Ref. 253 Can you deduce the approximate expression

$$\tan \theta_{\mathrm{u\,mix}} = \sqrt{\frac{m_{\mathrm{u}}}{m_{\mathrm{c}}}} \tag{204}$$

Challenge 213 r for the mixing of the up quark from the strand model?

CP violation in quarks

The CP violating phase δ for quarks is usually expressed with the *Jarlskog invariant*, defined as $J = \sin\theta_{12}\sin\theta_{13}\sin\theta_{23}^2\cos\theta_{12}\cos\theta_{13}\cos\theta_{23}\sin\delta$. This involved expression is independent of the definition of the phase angles and was discovered by Cecilia Jarlskog, an important Swedish particle physicist. Its measured value is $J = 3.06(21)\cdot 10^{-5}$. Ref. 233

Because the strand model predicts three quark generations, the quark model implies the possibility of CP violation. In the section on mesons we have seen that the strand model actually predicts the existence CP violation. In particular, Figure 99 shows that Page 338 Page 335

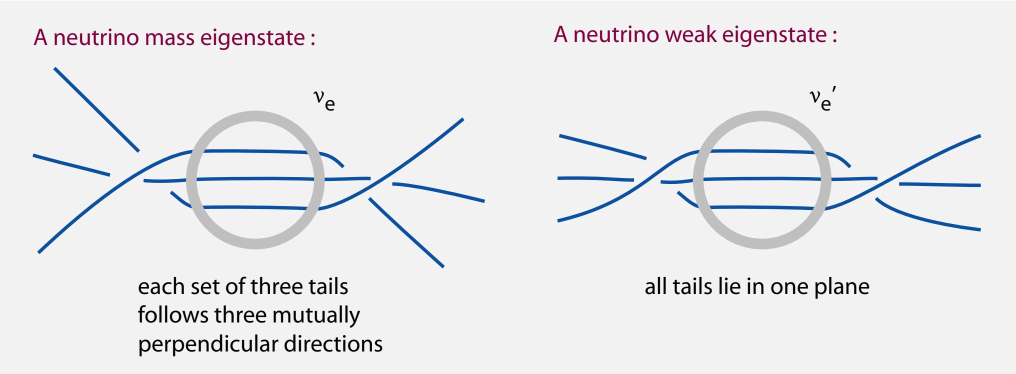


FIGURE 118 Tail shifting leads to neutrino mixing: mass eigenstates and weak eigenstates differ.

with the help of tail shifting, K^0 and $\overline{K}^0$ mesons mix, and that the same happens with certain other neutral mesons. Figure 100 shows a further example. As just mentioned, the possibility of tail shifting implies that CP violation is small, but non-negligible – as is observed.

The strand model thus predicts that the quark mixing matrix has a non-vanishing CP-violating phase. The value of this phase is predicted to follow from the geometry of the quark tangles, as soon as their shape fluctuations are properly accounted for. This topic is still a subject of research.

Neutrino mixing

The observation, in 1998, of neutrino mixing is comparably recent in the history of particle physics, even though the important physicist Bruno Pontecorvo predicted the Ref. 254 effect already in 1957. Again, the observation of neutrino mixing implies that also for neutrinos the mass eigenstates and the weak eigenstates differ. The values of the mixing matrix elements are only known with limited accuracy so far, because the extremely Ref. 233 small neutrino mass makes experiments very difficult. Experimental progress across the world is summarized on the website www.nu-fit.org. The absolute value of the so-called PMNS mixing matrix U is

$$|U| = \begin{pmatrix} 0.82(1) & 0.54(2) & -0.15(3) \\ -0.35(6) & 0.70(6) & 0.62(6) \\ 0.44(6) & -0.45(6) & 0.77(6) \end{pmatrix} . \qquad (205)$$

Again, these numbers are unexplained fundamental constants of nature. Within experimental errors, the matrix U is unitary. The mixing among the three neutrino states is strong, in contrast to the situation for quarks. Neutrino masses are known to be positive; however, present measurements are not precise and only yield values of the order of 1 ± 1 eV.

In the strand model, the lepton mass eigenstates correspond to tangles whose tails point along the three coordinate axes. In contrast, the weak eigenstates again correspond to tangles whose tails lie in a plane. The two kinds of eigenstates are illustrated in Figure 118. Again, the transition between the two eigenstates is due to tail shifting, a special kind of strand deformation.

We thus deduce that neutrino mixing, like quark mixing, is an automatic result of the strand model and is related to the weak interaction. Given that the neutrino masses are small and similar, and that neutrinos do not form composites, the strand model predicts that the mixing values are large. This is a direct consequence of tail shifting, which in the case of similar masses, mixes neutrino tangles leads to large mixings between *all* generations, and not only between neighbouring generations. In the strand model, the large degree of neutrino mixing is thus seen as a consequence of their low and similar masses, of their tangle structure, and of their existence as free particles.

Like for quarks, the strand model predicts a *unitary* mixing matrix for neutrinos. The strand model also predicts that the geometry of the neutrino tangles and their fluctuations will allow us to calculate the mixing angles. More precise predictions are still subject of research.

CP violation in neutrinos

The strand model predicts that the three neutrinos are massive Dirac particles, not Majorana particles. This has not yet been confirmed by experiment. The strand model thus predicts that the neutrino mixing matrix has *only one* CP-violating phase. (It would have three such phases if neutrinos were Majorana particles.) The value of this phase is predicted to follow from the neutrino tangles and a proper accounting of their fluctuations. Also this calculation is still a subject of research.

On the one hand, the strand model suggests the appearance of CP violation in neutrinos. On the other hand, it is unclear when the value of the CP-violating phase will ever be measured with sufficient precision. This is one of the hardest open challenge of experimental particle physics.

The mechanism of CP violation has important consequences in cosmology, in particular for the matter–antimatter asymmetry. Since the strand model predicts the absence of the see-saw mechanism, the strand model rules out leptogenesis, an idea invented to explain the lack of antimatter in the universe. The strand model is more on the line with electroweak baryogenesis.

Ref. 255
Ref. 256

Open challenge: calculate mixing angles and phases ab initio

Calculating the mixing angles and phases ab initio, using the statistical distribution of strand fluctuations, is possible in various ways. In particular, it is interesting to find the relation between the probability for a tail shift and for a tail braiding. This will allow checking the statements of this section.

Challenge 214 ny

Because the strand model predicts a lack of new physics beyond the standard model of particle physics, the calculation of neutrino mixing angles is one of the few possible *predictions* that are left over in fundamental physics. Since the lepton tangles are still tentative, a careful investigation is necessary.

One possibility is that only the electron neutrino tangle given above is correct, and that the other two neutrinos are similar to it, just with more built-in torsion. Figure 119 illustrates this possibility. If this assignment were correct, two of the mixing angles should be large (and maybe have the zero-order value of 120°/3=40°). In addition, very low mass

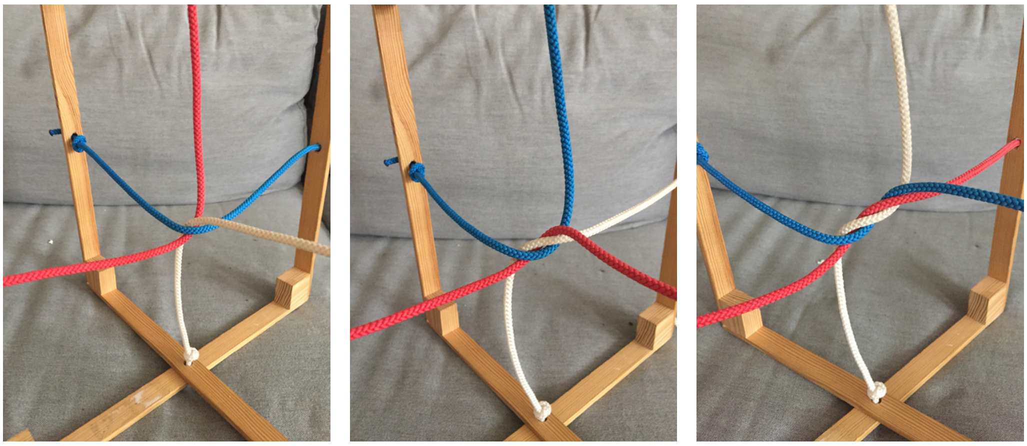


FIGURE 119 An alternative candidate assignment for the three neutrino tangles that generates large mixing between neighbouring generations and strong preference for one handedness.

values would arise naturally, in normal ordering. This tangle assignment would also explain the difficulty of observing neutrinos of opposite handedness: neutrinos would have an extremely high preference for one handedness (belt-trick); the mirror images of the tangles in Figure 119 would correspond to antineutrinos. Despite these appealing aspects, this tentative assignment has one unclear issue: explaining the lack of a fourth neutrino generation is not straightforward.

Summary on mixing angles and the millennium list

The strand model implies that mixing angles for quarks and neutrinos are properties of their tangle families. The existence of mixing is due to the shape of tangles and their fluctuations. As a result, strands explain why mixing angles are not free parameters, but discrete and unique constants of nature. The strand model also predicts that mixing angles are constant during the evolution of the universe.

We have shown that tangles of strands predict non-zero mixing angles for quarks and neutrinos, as well as CP-violation in both cases. The strand model also predicts that the mixing angles of quarks and neutrinos can be calculated from strand fluctuations. Strands predict that mixing matrices are unitary and that they run with energy. Strands also predict a specific sequence of magnitudes among matrix elements; the few predictions so far agree with the experimental data. Finally, the strand model rules out leptogenesis.

We have thus partly settled four further items from the millennium list of open issues. All qualitative aspects and some sequences are reproduced correctly, but no hard quantities were deduced yet. The result is somewhat disappointing, but it is also encouraging. At present, no other explanation for quark and neutrino mixing is known. Future calculations will allow either improving the checks or refuting the strand model. We leave this topic unfinished and proceed to the most interesting topic that is left: understanding the coupling constants.

Page 164

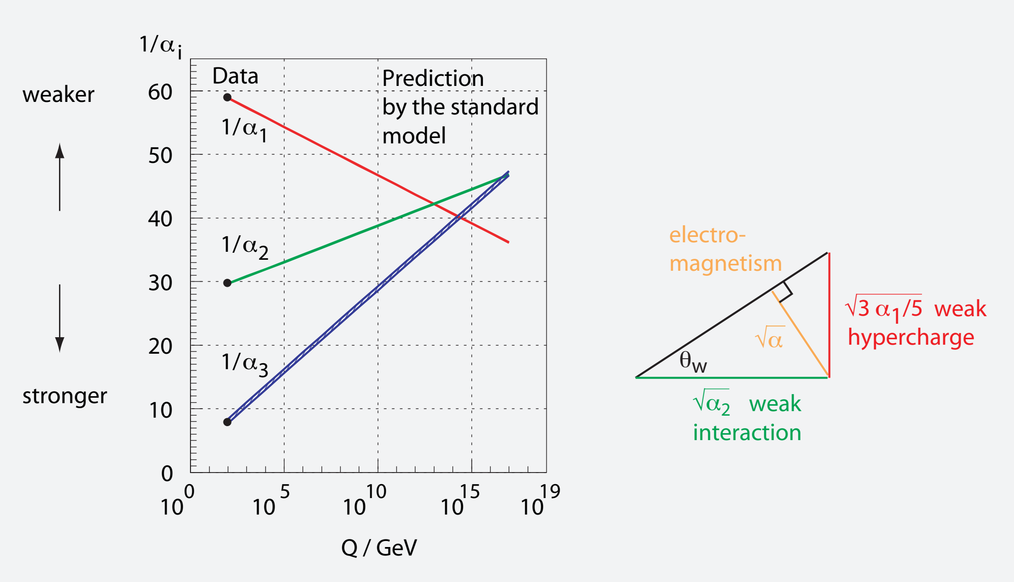


FIGURE 120 Left: How the three coupling constants (squared) change with energy, as predicted by the standard model of particle physics; the graph shows the constant $\alpha_1 = \frac{5}{3}\alpha/\cos^2\theta_W$ for the weak hypercharge coupling (related to the electromagnetic fine structure constant α through the weak mixing angle θ_W and a historical factor 5/3 that is useful in grand unification), the coupling constant $\alpha_2 = \alpha_w = \alpha/\sin^2\theta_W$ for the weak interaction, and the coupling constant $\alpha_3 = \alpha_s$ for the strong interaction. The three black points are measurement points; at lower and slightly higher energies, data and calculation match within experimental errors. (Courtesy Wim de Boer) Right: The relation between the coupling constants α for the electromagnetic $U(1)_{EM}$, $\alpha_2 = \alpha_w$ for the weak SU(2), α_1 for the weak hypercharge $U(1)_Y$ gauge groups and the weak mixing angle θ_W.

Coupling constants and unification

In nature, electric, weak and strong charge are *quantized*. No experiment has ever found even the smallest deviation from charge quantization. *All charges in nature are integer multiples of a smallest charge unit.* Specifically, the electric charge of every free particle is observed to be an integer multiple of the positron electric charge. We call the integer the *electric charge quantum number*.

In nature, the *strength* of a gauge interaction for a unit charge is described by its coupling constant. The coupling constant gives the probability with which a unit charge emits a virtual gauge boson, or, equivalently, the average phase change produced by the absorption of a gauge boson. There are three charge types and three coupling constants: for the electromagnetic, for the weak and for the strong interaction. All particles with a given charge type and value share the same coupling constant, even if their masses differ. The three coupling constants depend on energy. The known data and the change with energy predicted by the standard model of particle physics are shown in Figure 120.

In nature, the *fine structure constant* α, i.e., the electromagnetic coupling constant, at the lowest possible energy, 0.511 MeV, has the well-known measured value Ref. 5

$$\alpha = 1/137.035\,999\,139(31)\ . \qquad (206)$$

Equivalently, the electromagnetic coupling of the positron can also be described by the equivalent number

$$\sqrt{\alpha} = 1/11.706\,237\,6167(13) = 0.085424543114(10)\,, \qquad (207)$$

which is also called the *electric charge unit* (at low energy). Quantum electrodynamics predicts the precise change with energy of this charge unit; the experiments performed so far, up to over 100 GeV, agree with this prediction. Quantum electrodynamics also predicts that the charge unit, when extrapolated right up to the Planck energy, would have a value of 1/10.2(1). These predictions are shown, in a common, but somewhat scrambled way, in Figure 120.

Explaining the value of α, which determines all colours and all material properties in nature, is the most famous millennium issue. If the strand model cannot reproduce every observation about α and the other coupling constants, it is wrong. In particular, we thus need to understand, using the strand model, the quantization of charges on the one hand, and the origin of the mysterious value of the charge unit – either at low energy or at Planck energy – on the other hand.

Interaction strengths and strands

In the strand model, *all three gauge interactions are due to shape changes of tangle cores.* We first classify the possible shape changes. Given a tangle core, the following shape changes can occur:

— *Small* changes of core shape do not produce any crossing switch. Small shape changes thus have no physical significance: for a given observer, they leave all observables unchanged.
— *Twist* shape changes of a strand segment in the core produce an *electric field*, if the particle is charged. More precisely, the electric field around a particle is the difference between the average number p_{tr} of right twists and the average number p_{tl} of inverse, left twists that a particle tangle produces per unit time.
— *Poke* shape changes of a strand segment in the core produce a *weak interaction field*. More precisely, the weak field is the asymmetry among the probabilities $p_{\mathrm{p}x}$, $p_{\mathrm{p}y}$ and $p_{\mathrm{p}z}$ for the three fundamental poke types and their inverses.
— *Slide* shape changes of a strand segment in the core produce a *colour field*, if the particle has colour. More precisely, the colour field is the asymmetry among the probabilities $p_{\mathrm{s}1}$ to $p_{\mathrm{s}8}$ for the eight fundamental slide types and their inverses.
— A combination of these moves can also appear.

In the strand model, the fluctuation probabilities for each Reidemeister move – twist, poke or slide – determine the coupling constants. We thus need to determine these probability values. We can directly deduce a number of conclusions, without any detailed calculation:

— The coupling constants are not free parameters, but are specified by the geometric, *three-dimensional shape* of the particle tangles.
— By relating coupling constants to shape fluctuation probabilities, the strand model

predicts that coupling constants are *positive* numbers and *smaller than 1* for all energies.

▷ $\alpha < 1$,

▷ $\alpha_w < 1$,

▷ $\alpha_s < 1$

This is indeed observed.

— A still stricter bound for coupling constants can also be deduced. The sum of all possible fluctuations for a particular tangle has unit probability. We thus have

$$1 = p_{\text{small}} + p_{\text{tr}} + p_{\text{tl}} + \sum_{w=x,y,z} (p_{\text{p}w} + p_{\text{p}-w}) + \sum_{g=1}^{8} (p_{\text{s}g} + p_{\text{s}-g}) + p_{\text{combination}} . \quad (208)$$

The strand model thus predicts that the *sum* of the three charge units must be *strictly smaller than 1,* for every energy value. This is easily checked, both with the data and with the prediction of quantum field theory. In quantum field theory, the three (modified) coupling constants are given, as a function of energy, in the popular graph shown in Figure 120. The values are a combination of experimental data – for low energies – and theoretical extrapolations – for high energies. In this popular graph, the electromagnetic coupling is traditionally multiplied by $5/(3\cos^2\theta_W)$. (This is done in order to test grand unification; we keep the traditional factor, even though grand unification is shown by experiment and predicted by the strand model not to apply to nature.) The graph allows us to confirm that the sum of the three unmodified charge units is indeed smaller than 1 for all energy values, as predicted by the strand model.

— The strand model also predicts that the three coupling constants are related by small numbers, as the corresponding fluctuations differ only in the number of involved strands. This is also observed, as Figure 120 shows – especially if we remember that the couplings are the square roots of the values shown in the graph, corrected for the traditional factor.

— The strand model further predicts that the coupling constants are independent of time and space, and that in particular, they do not depend on the age of the universe. This is also observed, despite occasional claims to the contrary. Ref. 257

— Finally, strand model predicts that the coupling constants are the same for particles and antiparticles, as is observed.

In summary, the strand model implies, like quantum field theory, that *coupling constants are probabilities.* In addition, strands imply

▷ $\alpha + \alpha_w + \alpha_s < 1$ and $\sqrt{\alpha} + \sqrt{\alpha_w} + \sqrt{\alpha_s} < 1$.

Despite the agreement with experiment, we have not deduced any deep result yet – except one.

Strands imply unification

In fact, one new point is made by the strand model. Each gauge interaction is due to a different Reidemeister move. However, given a specific tangle core deformation, different observers will classify the deformation as a different Reidemeister move. Indeed, *every* Reidemeister move can be realized by the *same* deformation of a single strand: for each Reidemeister move, it is sufficient to add a curved section to a straight strand segment. Such a deformation can look like a type I Reidemeister move for one observer, like a type II move for another, and like a type III move for a third one.

Page 277

Because all interactions follow from the same kind of strand deformation of tangle cores, the strand model thus provides *unification* of the interactions. This result is new: in fact, this kind of strand unification of the interactions differs completely from any other approach ever proposed. And in contrast to several other approaches, strand unification does *not* require that the three coupling constants have the same value at high energy.

A given shape deformation thus has five probabilities associated to it: the probabilities describe what percentage of observers sees this deformation as a type I move, as a type II move, as a type III move, as a combination of such moves, or as no move at all, i.e., as a small move without any crossing switch. On the other hand, at energies measurable in the laboratory, the moves can *almost always* be distinguished, because for a given reaction, usually all probabilities but one practically vanish, due to the time averaging and spatial scales involved.* In short, at energies measurable in the laboratory, the three gauge interactions almost always differ.

Challenge 215 e

Calculating coupling constants

The strand model predicts that the calculation of the three coupling constants is a problem of tangle geometry and fluctuation statistics. Thus it can be approached, at each energy scale, either *analytically* or with *computer calculations*. The calculations need to determine the probabilities of the corresponding Reidemeister moves. If the results do not agree with the experimental values, the strand model is false. We note that there is no freedom to tweak the calculations towards the known experimental results.

In particular, in the strand model, one way to proceed is the following. The (square root of the) fine structure constant is the probability for the emission of twists by a fluctuating chiral tangle.

▷ The strand model predicts that the fine structure constant can be calculated by determining the probability of twists, i.e., Reidemeister I moves, in the *fluctuating tangle shapes* of a given particle with nonzero electrical charge.

In other words, the strand model must show that the probability of the first Reidemeister move in chiral particle tangles is *quantized*. This probability must be an integer multiple of a unit that is common to all tangles; and this coupling unit must be the fine structure

* The strand model thus appears to predict that at extremely high energy, meaning near the Planck energy, for each gauge interaction, also particles with zero charge can interact. At Planck energy, when horizons form, the time averaging is not perfect, and interactions become possible even with zero charge. But then the particle concept make little sense at those energies.

constant. Any check for the existence of a coupling unit requires the calculation of twist appearance probabilities for *each* chiral particle tangle. The strand model is only correct if all particles with the *same* electric charge yield the *same* twist emission probability.

Instead of emission, also absorption can be used to calculate the fine structure constant:

▷ The strand model predicts that the fine structure constant can be calculated from the average angle that a tangle core rotates when absorbing a photon.

We will pursue this alternative shortly.

So far, there do not seem to exist any analytical tool that permits the calculation of shape deformation probabilities. Thus, at present, computer calculations seem to be the only possible choice. Of all existing software programs, the most adapted to calculating fluctuation probabilities are the programs that simulate the dynamics of tangled polymers; but also the programs that simulate the dynamics of cosmic strings or the dynamics of helium vortices are candidates. The main issue, apart from a large computer time, is the correct and self-consistent specification of the shape fluctuation distribution at each energy scale.

Challenge 216 r

In summary, using the strand model we expect to be able to calculate the electromagnetic coupling constant and to understand its validity across all elementary particles. The same expectation obviously also holds for the two nuclear interactions. If any of the expectations on tangle interactions are found to be incorrect, the strand model is false. *The strand model must yield quantized tangle equivalence classes for the electromagnetic, weak and colour charge.* Even though the calculation issues are still subject of research, there are encouraging hints that these expectations will be validated.

First hint: the energy dependence of physical quantities

In nature, all effective charges, i.e., the coupling constants, change with energy. One also says that they *run* with energy. Figure 120 shows the details. Running also occurs for masses and mixing angles. All other intrinsic particle properties, such as spin, parities and all other quantum numbers, are found *not* to change with energy. For the coupling constants, the measured changes between everyday energy and about 100 GeV agree with the prediction from quantum field theory.*

Page 382

The strand model predicts

▷ Coupling constants, masses and mixing angles change with energy because they are quantities that *depend* on the average *geometrical details,* and in particular, on the *scale* of the underlying particle tangles.

* In the standard model of particle physics, the running of the electromagnetic and weak coupling constants – the slope in Figure 120 – depends on the number of existing Higgs boson types. The (corrected) strand model predicts that this number is one. Measuring the running of the constants thus allows checking the number of Higgs bosons. Unfortunately, the difference is small; for the electromagnetic coupling, the slope changes by around 2 % if the Higgs number changes by one. But in future, such a measurement accuracy might be possible.

Ref. 229
Page 330

More precisely, the running quantities depend on the fluctuations of the geometric tangle shapes, and these fluctuations depend somewhat on the spatial and thus the energy scale under consideration. We note that the strand model predicts a running *only* for these three types of observables; all the other observables – spin, parities or other quantum numbers – are predicted to depend on the *topology* of the particle tangles, and thus to be *independent* of energy. This prediction agrees with observation. Therefore, we can now explore the details of the running.

Second hint: the running of the coupling constants at low energy

The strand model proposes a new view on the screening and antiscreening effects that are part of quantum field theory. In the strand model, screening effects are consequences of the statistics of shape deformations for loose tangle cores that are embedded into the strands that form the vacuum. Since these statistical effects can in principle be calculated, it is expected that such calculations can be compared with the predictions of quantum field theory shown in Figure 120. This check is in progress. A few results, however, can be deduced without any calculations at all.

Page 382

In the strand model, the electromagnetic interaction is due to the first Reidemeister move, the twist. For a charged particle – thus one with a chiral tangle core – the average difference in the occurrence of right and left twists determines the effective charge. It is expected that this difference *decreases* when the strand core is loose, because the loose strands are more similar to those of the surrounding vacuum, so that the differences due to the chirality of the tangle will be washed out. In the language of quantum field theory, the virtual particle-antiparticle pairs – created by the fluctuations of the vacuum strands – screen the central, naked charge. The screening is reduced when the energy is increased, and thus when the scales are reduced. In other words, the strand model predicts that the electromagnetic coupling *increases with energy*, as is observed:

$$\triangleright \quad \frac{d\alpha}{dE} > 0 \,.$$

Also for the two nuclear interactions, the washing out effect for loose tangle cores by the vacuum does occur as predicted by quantum field theory. In the weak interaction, the *antiscreening* of the weak charge appears in this way. In the strong interaction, both virtual quark–antiquark pairs and virtual gluon pairs can appear from the strands that make up the vacuum. Virtual quark–antiquark pairs lead to screening, as virtual electron–antielectron pairs do for the electromagnetic interaction. In addition, however, we have seen that the strand model of mesons implies that virtual gluon pairs lead to antiscreening. (In contrast, virtual photon pairs do not lead to such an effect.) Because the strand model fixes the number of quark and gluons, the strand model is consistent with the result that the screening of the colour charge by quark pairs is overcompensated by the antiscreening of the virtual gluon pairs.

Page 332

In other words, the strand model reproduces the observed signs for the slopes of the coupling constants in Figure 120, for the same reason that it reproduces the quantum field theoretic description of the three gauge interactions. The predicted running could also be checked quantitatively, by taking statistical averages of tangle fluctuations of varying dimension. This is a challenge for future research.

Third hint: further predictions at low energy

As we just saw, the complete explanation of the running of the couplings depends on the explicit boson and fermion content of nature and on the fact that the strand model reproduces quantum field theory. Interestingly, the strand model also proposes a simpler, though less precise explanation of the running.

At energies much smaller than the Planck energy, such as everyday energies, the strand model implies that the average size of the tangle core is of the order of the position uncertainty of a particle. In other words, any thickness of the strands – real or effective – can be neglected at low energies. Therefore, at low energies, the average strand length within a particle tangle core is also of the order of the de Broglie wavelength. Low, everyday energy thus implies *large*, *loose* and *spherical/ellipsoidal* tangle cores.

At low energies, shape fluctuations can lead to any of the three Reidemeister moves. The probabilities of such shape deformations will scale with some power of the average strand length within the tangle core. In other words, coupling constants depend on energy. But how exactly?

We note directly that *higher* Reidemeister moves, which involve larger numbers of strand segments, will scale with *larger* power values. In particular, the longer the strand in the core – i.e., the lower the energy – the more the relative probability for the higher Reidemeister moves will increase.

In summary, the strand model predicts that when a tangle is loose and long, i.e., when energies are low, the strong nuclear interaction, due to the third Reidemeister move, is the strongest gauge interaction, followed by the weak nuclear interaction, due to the second Reidemeister move, in turn followed by the electromagnetic interaction:

$$\triangleright \quad \alpha_{\mathrm{em}} < \alpha_{\mathrm{w}} < \alpha_{\mathrm{s}} \,.$$

The prediction matches observations. Unfortunately, this argument is not reliable. If the strand number were the *only* cause of the running, the argument would imply that the three slopes for the running of the three coupling constants should behave like 3:2:1. Page 382 However, the graph of Figure 120 shows otherwise, even if the difference between the electromagnetic coupling and the weak hypercharge coupling is taken into account. Indeed, the running of the coupling constants is not due to strand number only, but also to the explicit boson and fermion content of nature, as we just saw.

The running of the coupling constants up to Planck energy

At energies near the Planck energy, quantum field theory is modified: effects due to the strand diameter start to play a role. Near Planck energy, tangles get tighter and tighter and fluctuations get weaker, because there is less room for them. In other words, near Planck energy tangles tend to approach the structure of horizons. Therefore, near the Planck energy, the strand model predicts deviations from the energy dependence of the coupling constants that is predicted by quantum field theory. So far, estimating such deviations has not been possible.

Another calculation might seem more promising: to calculate the coupling constants near Planck energy. It could be argued that the approach to calculate the low-energy coupling constants from Planck-energy values seems unsatisfactory, due to the approximations and extrapolations involved. But it is possible if we are convinced that quantum field theory is correct up to Planck energy. And this is just what the strand model predicts. Such a Planck-scale calculation might then allow us to estimate the low-energy coupling constants from their Planck energy values. However, so far, also this approach has not led to success, despite a number of attempts. The challenge seems to be to understand core deformation for case of tight tangle cores. We keep this option in mind.

Limits for the fine structure constant do not provide explanations

When searching for ways to determine the fine structure constant, we need to be careful. Here is an example that explains why.

Numerous observations of nature imply a limit on the fine structure constant. A pretty one appeared in a post on the internet in 2017. The *electrostatic repulsion* between two electrons at a given distance must be larger than the *radiation force* between to small neutral black holes at that same distance. In other words,

Ref. 258

$$\frac{e^2}{4\pi\varepsilon_0}\,\frac{1}{r^2} > \frac{L_{\mathrm{bh}}}{c}\,\frac{\pi R_{\mathrm{bh}}^2}{4\pi r^2}\,. \tag{209}$$

Here it is assumed that thermal radiation from one black hole acts on the cross section of the other black hole by pushing it away. Multiplying both sides by $r^2/\hbar c$ and inserting the expressions for the black hole luminosity L_{bh} and the black hole radius R_{bh} gives

Vol. V, page 151

$$\alpha > \frac{1}{15\,320\,\pi}\,. \tag{210}$$

The bound is not tight, but is obviously correct.

Various researchers are looking for observations that give the best possible bound for α. Such a search can indeed yield much better bounds. However, such a search cannot *explain* the value of α. We can indeed use thermodynamics, gravity or other observed properties to deduce *observational limits* on α. Many formulae of physics contain α in a more or less obvious way. Maybe, one day, known physics will be able to yield very tight upper and lower bounds for α. Still, *the explanation of the value of α would still lack.*

Ref. 259

To explain the fine structure constant α, we need an approach based on the complete theory, not one based on known, millennium physics, such as expression (209). Millennium physics can *measure* α, but *cannot explain* it. To explain the fine structure constant, a unified theory is needed. In our case, we need to check whether we can calculate α with strands. Therefore, we now explore tangle topology, tangle shapes and tangle motion with this aim in mind.

Charge quantization and topological writhe

In the strand model, electric charge is related to the *chirality* of a tangle. Only chiral tangles are electrically charged. The strand model thus implies that a topological quantity for tangles – defined for each tangle in the tangle family corresponding to a specific elementary particle – must represent electric charge. Which quantity could this be?

The first candidate for charge in the strand model is provided by knot theory:

▷ The usual topological quantity to determine chirality of knots and tangles is the *topological writhe*.

To determine its value, we draw a *minimal projection*, i.e., a two-dimensional knot or tangle diagram with the *smallest* number of crossings possible. We then count the right-handed crossings and subtract the number of left-handed crossings. This difference, an integer, is the topological writhe. Topological writhe is thus a two-dimensional concept and does not depend on the shape of a knot or tangle. We note:

— The topological writhe of the W boson tangles is +3 or −3, depending on which mirror image we look at; the topological writhe of the Z boson and Higgs boson tangles vanishes. The topological writhe of any unknotted strand also vanishes. In this way, if we define the electric charge quantum number as *one third* of the topological writhe, we recover the correct electric charge quantum number of the weak and all other gauge bosons. We note that the Higgs boson does not change this result, so that all family members of a particle share the same topological writhe.

Page 320 — The tangles of the quarks show that if we define the electric charge quantum number as *one third* of the topological writhe, we recover the correct electric charge quantum number of all quarks. Adding Higgs bosons has no effect on this definition.

Page 326 — The tangles of the leptons show that if we define the electric charge quantum number as the topological writhe of the *centre region* only, we recover the correct electric charge quantum number of all leptons. Again, adding Higgs bosons does not change this result.

In other terms, the electric charge quantum number can be reproduced with the help of topological writhe. And indeed, the electric charge of massless bosons, i.e., photons and gravitons, vanishes.

Let us sum up. In nature, electric charge is quantized. The strand model describes charged particles with the help of fluctuating alternating tangles, and charge quantization is a topological effect that results because all particles are made of strands. In particular,

▷ The *electric charge quantum number* behaves similarly to topological writhe (times one third or times one): it is quantized, has two possible signs, vanishes for achiral tangles, is a topological invariant – and thus is conserved.

In short, a topological quantity, namely topological writhe, reproduces the electric charge quantum number in the strand model. Three issues remain. First, given that every particle is described by a tangle family with an infinite number of members, how is Challenge 217 e the electric charge, i.e., the topological writhe of the other tangle family members accounted for? It is not hard to see that family members do not change topological writhe. The second issue is more thorny: why is the charge definition different for leptons? We skip this problem for the time being. The third issue is the central one: What is the origin of the peculiar value of the charge unit, whose square has the value $\alpha = 1/137.035\,999\,139(31)$ at low energy?

Challenge 218 ny
Ref. 5

Charge quantization and linking number

An alternative conjecture for charge quantization is the following:

▷ Electric charge, i.e., twist emission probability, might be proportional to the *linking number* of ribbons formed by strand pairs.

The following arguments speak in favour of this conjecture.

— In knot theory, a ribbon is the strip associated to and limited by two strands.
— The *linking number* of a ribbon is the number of times that the two edges of a ribbon wind around each other. The linking number is a topological invariant and an integer.
— In particle tangles, only wound up, i.e., linked ribbons should lead to (net) boson emission. For tangles made of three strands, we define a total linking number as the sum of all three possible linking numbers.
— The linking number of the Higgs boson strand pairs is zero; that of the Z boson strand pairs is the sum of 1, 0 and -1, thus also zero. The linking number for the W boson is 3 or -3, that of the quarks is 1, -1, 2 or -2. We thus conjecture that the charge quantum number is one third of the total linking number.
— Massless bosons, i.e., photons, gluons and gravitons, have no electric charge.

In short, linking number, an integer, might be a better topological quantity to explain electric charge quantization than topological writhe. On the other hand, it might well be that linking number, being a quantity that depends on *two* strands, is related to the *weak charge* rather than to the electric charge.

If the conjectured relation between linking number and electric or weak charge is correct, it might lead to a calculation of the corresponding coupling constant, once the tangle shape or, better, once the tangle dynamics is included in the proper way. For example, the photon emission probability could depend on the *writhe* or on the *twist* of the (averaged) ribbons. Both these properties might lead to virtual photon emission. (The sum of writhe and twist of a ribbon is given by the linking number, as explained by Calugareanu's theorem.)

In this and any topological definition of electric charge, we face two slight hurdles: First, we have to watch out for the graviton: it is uncharged. Secondly, we have to explain why the strand model for the simplest family member of the d quark is not chiral. Both hurdles can be overcome.

Challenge 219 e

If the linking of *two* strands is connected to weak charge, it might well be that a similar quantity defined for *three* strands is related to colour charge. All these possibilities are topic of research.

How to calculate coupling constants

The strand model suggests that crossing number and linking number somehow define electric and weak charge. In simple words, the model suggests that quantization of all charge types is a topological effect; quantization is due to the multiple ways in which strands cross inside tangles.

Coupling constants describe the probability of interaction with gauge bosons. Experiments show that these quantities are slightly *scale*-dependent, since they run with energy. But in the strand model, coupling constants are not really *shape*-dependent: electrons, muons and antiprotons have the same electric charge and fine structure constant values despite being described by different tangles. Coupling constants do not depend on the *kind* of tangle. Experiments show that they just depend somewhat on its size. In short,

▷ We need a definition of each coupling constant that is *tangle-independent* and *shape-independent*, and only depends on a topological invariant of tangles.

In fact, this conclusion eliminates many speculations, including a number of calculation approaches that were included in this chapter in previous editions. We are left with just a few options. To explore them, we start with an overview.

Coupling constants in the strand model

In experiments, there are the following gauge interactions with their charges:
1. The electromagnetic interaction with electric charge and U(1) symmetry.
2. The weak interaction with weak isospin and SU(2) symmetry.
3. The strong interaction with colour and SU(3) symmetry.

In the strand model, the *gauge interactions* are modelled as transfers of Reidemeister moves:
1. The electromagnetic interaction is twist transfer and the electric charge is preferred twist transfer to or from a massive particle. Twists can be added abd form a circle: they form a U(1) Lie group. They change the tangle phase by exchanging one observable crossing.
2. The weak interaction is poke transfer and the weak isospin is preferred poke transfer to or from a massive particle. Pokes exist in three linearly independent directions and their generators behave like the belt trick: they generate an SU(2) Lie group. They change the tangle phase by exchanging two observable crossings.
3. The strong interaction is slide transfer and the colour charge is preferred slide transfer to or from a massive particle. Slides can be added, its generators have a Z_3 symmetry and they form an SU(3) Lie group. They change the tangle phase by exchanging two or three? crossings.

In the strand model, *neutral particles* are those that cannot receive Reidemeister moves or that receive them all in equal way:
1. Electromagnetism: Neutral 'tangles' are made of one strand (e.g., the photon) or are topologically achiral (the Z and the neutrinos).

2. Weak interaction: Neutral tangles are made of one strand (e.g., the photon) or of two straight or unpokeable strand pairs (the Z, the right-handed leptons and quarks).
3. Strong interaction: Neutral tangles are made of one strand or of three strands.

In the strand model, *charged particles* are specific tangles:
1. Electric charge is due to the observability of crossings during photon emission or absorption, i.e., when twists are applied. Particles with electric charge, i.e., with preferred twist transfer, have a global asymmetry, global twistedness, namely topological chirality. Locally, electrically charged particles have crossings; electric charge is positive or negative. Charge is 1/3 of the signed crossing number. Examples are the charged leptons, the quarks and the W boson.
2. Weak charge is thus due to the observability of crossings during W or Z emission or absorption, i.e., when pokes are applied. Particles with weak isospin, i.e., with preferred poke transfer, have a global asymmetry that prevents all pokes to act equally effectively: For fermions, such an asymmetry arises when tangle twistedness and the belt trick have the same sign; thus all left-handed fermions and right-handed antifermions have weak isospin. Locally, weakly charged fermions behave like a belt buckle that rotates in the appropriate direction. Due to their tangle topology, some fermions have positive, others negative weak isospin. For the W boson, the asymmetry is built into the tangle; due to the tangle structure, the W and its antiparticle have plus or minus twice the weak isospin of fermions.
3. Colour, strong charge, is due to the observability of crossings during gluon emission or absorption, i.e., when slides are applied. Particles with colour charge, i.e., with preferred slide transfer, have a global asymmetry that prevents all slides to act equally effectively: Coloured particles are made of exactly two strands with tails in tetrahedron skeleton directions. Only two-stranded tangles allow certain slides and prevent others. Therefore only quarks have colour charge. Locally, red, blue and green colours correspond to three directions in one plane that differ by an angle of $2\pi/3$.

Coupling strength is the ease of crossing rotation, of poke creation, and of slide induction. These connections allow calculating the coupling strength values.

Deducing α from precession

In nature, magnetic fields rotate charged particles. In the strand model, as shown in Figure 53, magnetic fields are made of moving twists. In fact, from the strand definition of the electromagnetic interaction and the electric charge and from the drawing in Figure 50, we deduce:

Page 234
Page 229

▷ Moving twists rotate crossings.

We note that this description differs slightly from a pure twist transfer. But this formulation is the key to calculating α.

We assume that the typical, average crossing is lying in the paper plane, as in the drawing of the fundamental principle. For an average crossing, the two strands lie along the x and y axes. When a photon, i.e., a twist, arrives along the diagonal in the first quadrant, it rotates the crossing completely, by one turn. If the twist arrives at a different angle, its effect is lowered. We approximate this angle effect with simple trigonometry: we assume that the angular projection describes the reduction of the effect with the incoming angle of the twist.

For the incident photon, we call γ the angle from the y-axis and β the angle out of the paper plane. The average rotation angle induced by an absorbed photon or twist on a charged particle with *three* crossings, corresponding to *one* elementary charge, can be calculated. We include $\sin\gamma$ for the volume element in spherical coordinates and average over the possible angle values δ between the strands at the crossing. Further terms arise from the trigonometric approximation. In particular, a second power arises from the two tails, and a further squaring is required to get probabilities. Nevertheless, the expression remains open to dispute:

$$\sqrt{\alpha_{\text{calc}}} = \frac{3}{2\pi^2} \int_{\delta=0}^{\pi} \int_{\gamma=-\delta/2}^{\delta/2} \int_{\beta=-\pi/2}^{\pi/2} \cos\beta\,(\sin\delta)^2\,(\cos(\gamma\pi/\delta)\,\cos\beta)^4\,\mathrm{d}\beta\,\mathrm{d}\gamma\,\mathrm{d}\delta = 0.15\,. \tag{211}$$

The resulting value of 0.15 is not an acceptable approximation to reality, in which $\sqrt{\alpha} = 0.08542454311(1)$ at low energy and $\sqrt{\alpha} = 0.10(1)$ at Planck energy. Neither is the value a good approximation to the hypercharge coupling, which changes from $\sqrt{\alpha_1} = 0.10(1)$ at 100 GeV to $\sqrt{\alpha_1} = 0.13(1)$ at Planck energy. We need a better approximation for the value of the electromagnetic coupling strength.

Deducing the weak coupling

Weak fields deform strand (crossing) pairs by adding or transferring generalized pokes. Weak fields are collections of pokes; pokes represent virtual weak bosons. The weak isospin, the weak charge, is related to the orientation of the strand pairs. The weak interaction occurs through an incoming poke that deforms a strand pair:

▷ A moving poke rotates a pair of strands.

This process is the key to calculating α_{w}. We note that there is a certain similarity to the setting used for calculating the electromagnetic coupling: in both cases, the incoming boson acts on a target consisting of two strands. This similarity is the reason for electroweak mixing.

We calculate the coupling constant for a single belt buckle, assuming parallel strands. The average rotation angle induced by one incoming weak (unbroken) boson (out of the three possible cases) is one full turn when the impact is perpendicular to the two strands and to the plane defined by them. For a general incidence angle the induced rotation angle is lower. We again use trigonometrical projection to approximate the induced crossing rotation angle in the general case, with the same issues as in the previous case. We call γ the angle from ideal incidence, and β the longitude. The average angle is then given by

$$\sqrt{\alpha_{\text{w calc}}} = \int_{\gamma=0}^{\pi/2} \int_{\beta=0}^{2\pi} \sin\gamma\,(\cos^2\gamma\,\cos^2\beta)^4\,\mathrm{d}\gamma\,\mathrm{d}\beta \approx 0.19\,. \tag{212}$$

If we need to average over the different angles between the strands that make up the pair experiencing the poke, we get a different value.

The calculated value of the weak coupling is not an acceptable approximation to reality, in which $\sqrt{\alpha_w} = 0.18$ at the (low) energy of 100 GeV and $\sqrt{\alpha_w} = 0.14$ at Planck energy. We need a better approximation.

Deducing the strong coupling

Strong fields deform specific three-strand configurations by adding generalized slides. The generalized slides are due to gluons. Strong colour is related to the order and orientation of the strands in these specific three-strand configurations. In short:

▷ Incoming, moving slides deform three-strand configurations.

This is the key to calculating α_s.

We assume that one of eight possible gluons is incident. In an average triple strand configuration, the three strands are oriented in a way that in the paper plane they look like three symmetrically arranged rays. One ray lies along the y axis. When a gluon arrives, it performs a slide. For an incident gluon, we call γ the angle from the y-axis to the next strand and β the angle out of the paper. In the trigonometric approximation, the average slide angle induced on a coloured particle is given by

$$\sqrt{\alpha_{\text{s calc}}} = \int_{\gamma=0}^{\pi/2} \int_{\beta=0}^{2\pi} \sin\gamma \, (\cos^3\gamma \; \cos^3\beta)^4 \, d\beta \, d\gamma \approx 0.11 \; . \tag{213}$$

This is not an acceptable approximation to reality, in which $\sqrt{\alpha_s} = 0.7(1)$ at the (low) energy of 1 GeV and $\sqrt{\alpha_s} = 0.13(1)$ at Planck energy. We need a better approximation for the strong coupling.

Open challenge: calculate coupling constants with precision

The approximations used above for estimating the coupling constants can be dismissed as mere educated guesses. Despite this objection, these guesses show that a determination of the coupling constants from the strand model is within reach, and that it can be realized with limited effort. It is sufficient to improve the three approximations; this is can be realized by using computer simulations for the transfer of Reidemeister moves or by finding an improved analytical model.

Calculating all three coupling constants ab initio with high precision will allow checking the statements of this section in an independent manner and, above all, will allow testing the strand model. The calculations should be performed at different energies, to confirm the energy dependence of the couplings. Also the influence of the effective strand diameter on the fine structure constant should be explored.

Challenge 220 r

In order to reach highest precision, the effects of the various tangle family members might have to be taken into account, because in the strand model, each particle is described by a family of tangles. On the other hand, the strand model predicts that family members have a small effect on the coupling constant, so that the family issue can be

Page 352

neglected in the beginning.

In the case of the nuclear coupling constants, Arnold's results on plane curves may help in the estimations and calculations. Ref. 260

Electric dipole moments

Experimental physicists are searching for electric dipole moments of elementary particles. No non-zero value has been detected yet. The idea of electric dipole moment is based on a non-spherical distribution of electric charge in space.

In the strand model, particles are tangles. As a consequence, the electric charge distribution – the distribution of the crossings in a tangle – is intrinsically a slightly non-spherical quantity, thus a quantity unequally distributed in space. However, it is only non-local on a scale of the order of a Planck length. In other terms, the electric dipole moment d of elementary particles is predicted to be

▷ $d = f\, e\, l_{\mathrm{Pl}}$,

where the factor f arises from averaging the tangle and is of order one. Similar values are predicted by the standard model in the absence of supersymmetry and grand unification. Ref. 261 However, the sensitivity of measurements has not reached these values yet, by several orders of magnitude.

We note that the strand model predicts that the dipole moment changes, or 'runs', with energy. This follows from the shape-dependence of the dipole moment. Such a dependence is also predicted by quantum field theory.

In summary, we expect that up to a region close to a Planck length, the *strand* model should not yield dipole moments that differ in order of magnitude from those predicted by the *standard* model of particle physics. In the future, more precise calculations and measurements could allow testing the strand model using dipole moments.*

Five key challenges about coupling strengths

There are many ways to evaluate candidates for unified models. A concrete evaluation focuses on four key challenges about coupling constants. These challenges must be resolved by any candidate model in order to be of interest.

1. So far, we explained particle charges with topological properties of the tangle models of the particles, and we explained coupling strengths with the transfer of crossings, pokes and slides. This allowed deducing a rough approximation of coupling constants. By doing so, we have settled a first key challenge:

▷ The strand model explains why the fine structure constant, or equivalently, the electric charge, is the *same* for electrons and protons.

Deducing this equality is a key challenge for any unified model. In fact, all coupling constants must be independent of particle type. This is indeed the case in the strand

* It might even be that software packages allowing to do this already exist. The package of Christian Beck that he used in arxiv.org/abs/hep-th/0702082 might be an example. Others might also exist.

model.

2. The second key challenge was the energy-dependence of the coupling constants. The strand model predicts that coupling constants run with energy in exactly the way that is predicted by QED, QCD and electroweak theory. We could also argue that this is not a real challenge for any unified model that reproduces these theories. In the strand model, the running of the electromagnetic coupling constant can be seen as a consequence of the gradual tightening of tangles with energy. For a typical electrically charged particle at low energy, the tangle is very loose; therefore:

▷ The Planck scale number of crossings is shielded by an additional cloud of crossings created by the loose strands of the tangle.

In this way, the strand model explains the running of the fine structure constant in exactly the same way as QED.

3. The third key challenge has only been touched upon very briefly:

▷ Any unified model needs to clarify the relation between the hypercharge, the electric charge and the weak isospin (the 'weak charge').

The strand model explains electromagnetism as acting on crossings and the weak interaction as acting on parallel strands. This general statement contains the required explanation; but the details still need to be worked out. It is expected that in electromagnetism, a *single* crossing is rotated, mainly by rotating *one* strand around the other. In contrast, in the weak interaction, *two* strands are rotated together, producing ar switching *two* crossings. The number of crossings differs between electromagnetism and the weak interaction, but the total number of involved strands is tow in both cases. As a result of this similarity, the two interactions mix. The final explanation of electroweak mixing might even allow to deduce a intuitive geometric meaning of θ_w , the weak mixing angle or Weinberg angle.

4. The fourth key challenge, related to the previous one, still needs to be explored in more detail:

▷ Any unified model must explain why the mass ratio of the intermediate weak vector bosons is related to the coupling ratio of the weak and the electromagnetic interaction as

$$\left(\frac{m_W}{m_Z}\right)^2 + \frac{\alpha}{\alpha_w} = 1 . \qquad (214)$$

The strand model strongly suggests that it can explain the relation, but the detailed argument must yet be provided. Using more drastic language, we can repeat what many have said already in the past: explaining the electroweak mixing expression (214) is the key challenge for any unified model.

In the strand model, the two electroweak coupling constants are measures for interaction probabilities of crossings with twists and with pokes. In contrast, masses are interaction probabilities of crossings with spatial curvature. Why are they related by expression (214)? Here is a short *brainstorm* on the issue.

In the strand model, mass appears by tail braiding. Tail braiding adds crossings, and in this way adds mass. Added crossings also imply added weak and sometimes electric charges. The Z boson arises from vacuum by different tail braidings than the W. The W arises by the braiding of two tail pairs at 90 degrees; the Z arises by braiding one tail pair at 90 degrees.

In case of the W and the Z bosons, the Z tangle produces a larger disturbance of the vacuum than the W; therefore it is more massive than the W.

At which angle does a clasp start to form a ''enclosed space in between''? Surprisingly, this happens for any angled strand pair, even if the angle is near π. How does this enclosed space change with scale, given that scale might change the clasp angle? This question might be related to the running of masses, mixing angles or coupling constants. In particular, we should answer the following question: Which physical observable does this enclosed space influence? Mass, couplings, or mixings? Is mass more related to ropelength or more related to the enclosed space?

5. The fifth key challenge is, of course, the precise calculation of the coupling constants.

Summary on coupling constants

The strand model implies that coupling constants are geometric properties of tangle families that correspond to charged particles. As a consequence, strands explain why the coupling constants are not free parameters in nature, but fixed constants. Strands predict that coupling constants are the same for particles with the same charge, and that coupling constants are constant during the macroscopic evolution of the universe. Strands predict small electric dipole moments for elementary particles, compatible with and lower than present measurement limits. Strands also predict the correct sequence of the coupling constants at low energy and the correct sign of their running with energy. Strands thus reproduce all observed qualitative properties of coupling constants. No other unified model achieves this yet.

Using tangle shapes, the strand model proposes several ways to calculate coupling constants ab initio. First estimates based on the new *tangled* particle models yield results of the order of the observed ones; nevertheless, the errors due to the approximations
Ref. 262 are still larger than the measurement errors. In 2019, a first publication has appeared . Improved calculations are ongoing and will allow either to confirm or to refute the strand model.

Chapter 13

A PICTORIAL SUMMARY OF THE STRAND MODEL

> La forma universal di questo nodo
> credo ch'i' vidi, ... **
>
> Dante, *La Divina Commedia*, Paradiso, XXXIII, 91-92.

Deducing all of modern physics from simple pictures is possible! Indeed, describing observations with the help of fluctuating strands agrees with all data and all experiments – without exception. More precisely, in the *strand model* or *strand conjecture*, all properties of nature follow by assigning the Planck units to fundamental events. This fundamental principle is illustrated in Figure 121: every fundamental event is a strand crossing switch occurring at the Planck scale. With this description, the strand model visualizes and explains quantum theory, the standard model of particle physics, general relativity and cosmology.

Quantum theory, inclusive wave functions and spin 1/2, is illustrated in Figure 122, Figure 123 and Figure 124. Gauge interactions and the origin of gauge couplings are illustrated in Figure 125 and Figure 126. The strand model proposes a tangle structure for each elementary particle; they are given in Figure 127 and Figure 128. The tangles determine masses and mixing angles. No modification or extension of the particle spectrum arises. With all the assigned particle tangles, Figure 129 and Figure 130 show that the strand model reproduces all Feynman diagrams of the standard model. No additional Feynman diagram is possible. In short: *No modification or extension of the standard model arises.*

The strand model also yields a tangle structure for flat vacuum, curved vacuum, gravitons, everyday gravity, and black hole horizons. They are shown in Figure 131, Figure 132, Figure 133, Figure 134 and Figure 135. These figures imply that *strands reproduce gravity and general relativity, without extension or modification at sub-galactic scales.*

Finally, the strand model proposes a tangle structure for the universe, from the origin to the present, as illustrated in Figure 136 and Figure 137. Strands predict a single universe with an expanding horizon, trivial topology and a matter density close to the critical value. The cosmological implications are still subject of research.

Strands fluctuating at the Planck scale thus explain the colours of flowers, the motion of butterflies, and the sight of the sky at night.

** 'The universal form of that knot, I think I saw, ...'

The **origin** of the strand conjecture

1. Planck units are limit values
$W \geqslant \hbar$
$F \leqslant c^4/4G$
$v \leqslant c$
$S \geqslant k$

2. Tethers explain spin 1/2

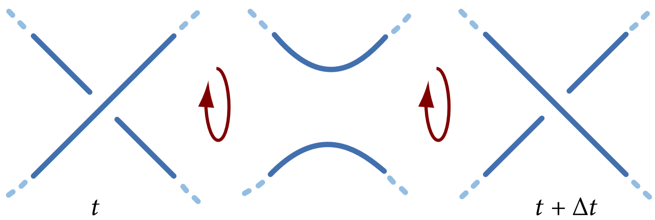
3. Extension explains black hole entropy and space microstructure

4. *Ur* or *qubit* as fundamental constituent

The **fundamental Planck-scale principle** of the strand conjecture

Strand conjecture

Observation



$W = \hbar/2$
$\Delta l \geqslant 2\sqrt{\hbar G/c^3}$
$\Delta t \geqslant 2\sqrt{\hbar G/c^5}$
$S = k/2$

A fundamental event localized in space

Predictions:

Events, observations and measurements are due to crossing switches.

All measurements are electromagnetic.

FIGURE 121 Many arguments lead to a description of nature with extended constituents at the Planck scale. The fundamental principle of the strand conjecture defines the simplest observation in nature, the almost point-like fundamental event. Every event results from a *skew strand crossing switch*, at a given position in three-dimensional space. The strands themselves are not observable; they are impenetrable and best imagined with Planck radius. The crossing switch all fundamental constants. The double Planck length limit and the double Planck time limit arise, respectively, from the smallest and from the fastest crossing switch possible.

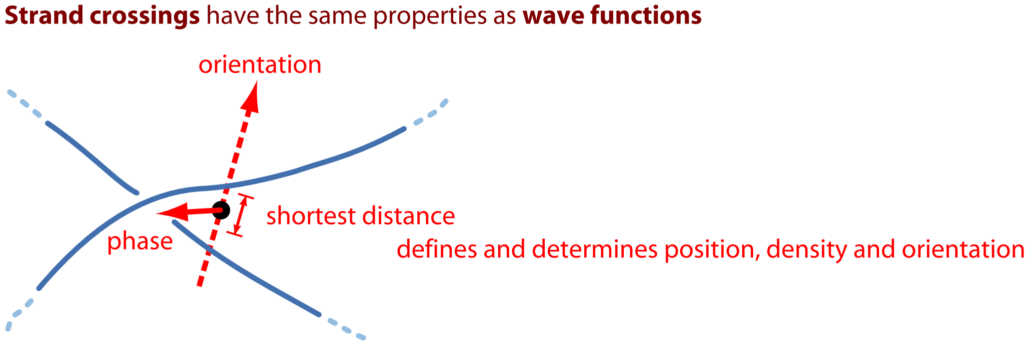


FIGURE 122 The geometry of a (skew) strand crossing suggests a relation to wave functions. In both cases, absolute phase around the orientation axis can be chosen freely. In contrast, phase differences due to rotations around that axis are always uniquely defined.

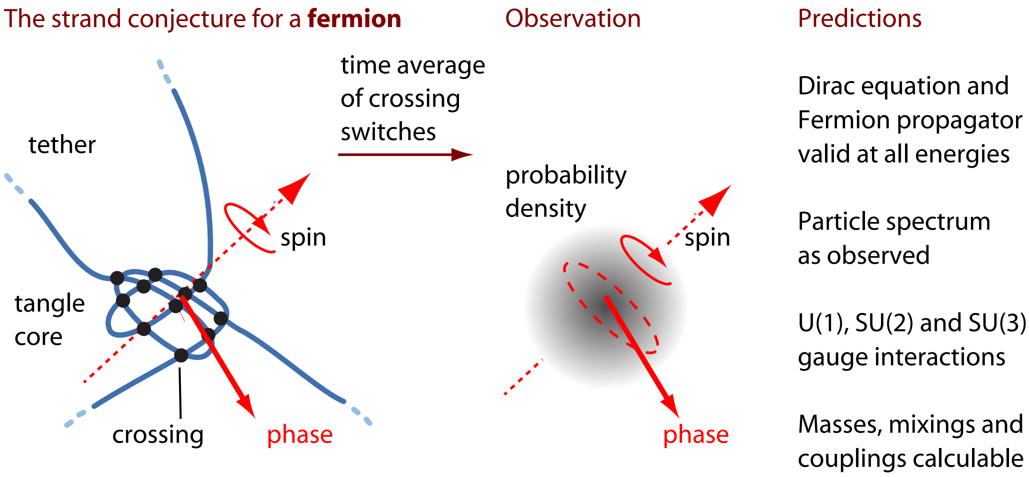


FIGURE 123 In the strand conjecture, the wave function and the probability density are due, respectively, to crossings and to crossing switches at the Planck scale. The wave function arises as time average of crossings in fluctuating tangled strands; a Hilbert space also arises. The probability density arises as time average of the crossing switches in a tangle. The tethers – connections that continue up to large spatial distances – generate spin 1/2 behaviour under rotations and fermion behaviour under particle exchange. The tangle model ensures that fermions are massive and move slower than light.

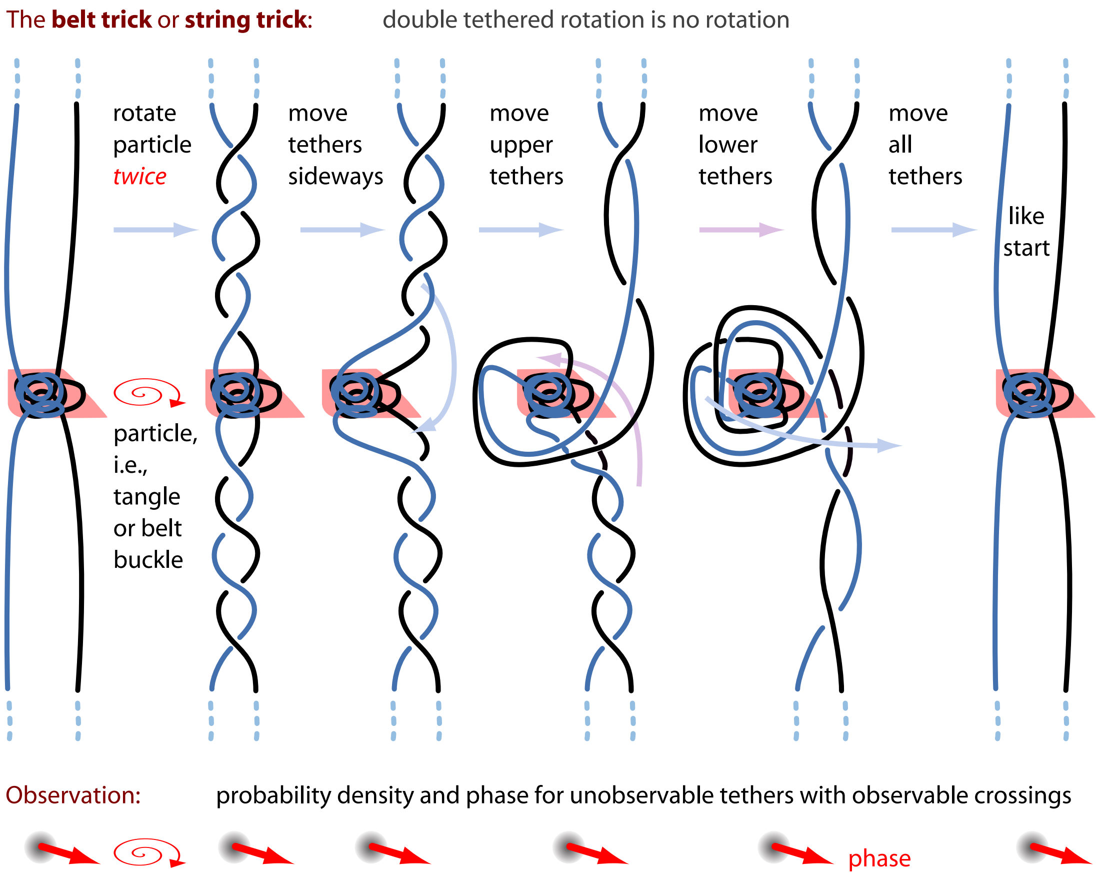


FIGURE 124 The *belt trick* or *string trick*: a rotation by 4π of a tethered particle, such as a belt buckle or a tangle, is equivalent to no rotation – when the tethers are allowed to fluctuate and untangle as shown. This equivalence allows the tethered particle to rotate forever. Untangling is impossible after a rotation by 2π only. This illustrates *spin 1/2*.
In addition (not shown), when two tethered particles are interchanged twice, all tethers can be untangled. Untangling is impossible after a single interchange only. This illustrates *fermion statistics*.
Both equivalences work for any number of tethers and assume that tethers are not observable, but crossing switches are.
The probability for the spontaneous occurrence of the belt trick depends on the complexity and details of the tangle core. This spontaneous process, together with Higgs emission and absorption, leads to the *mass* of elementary particles.

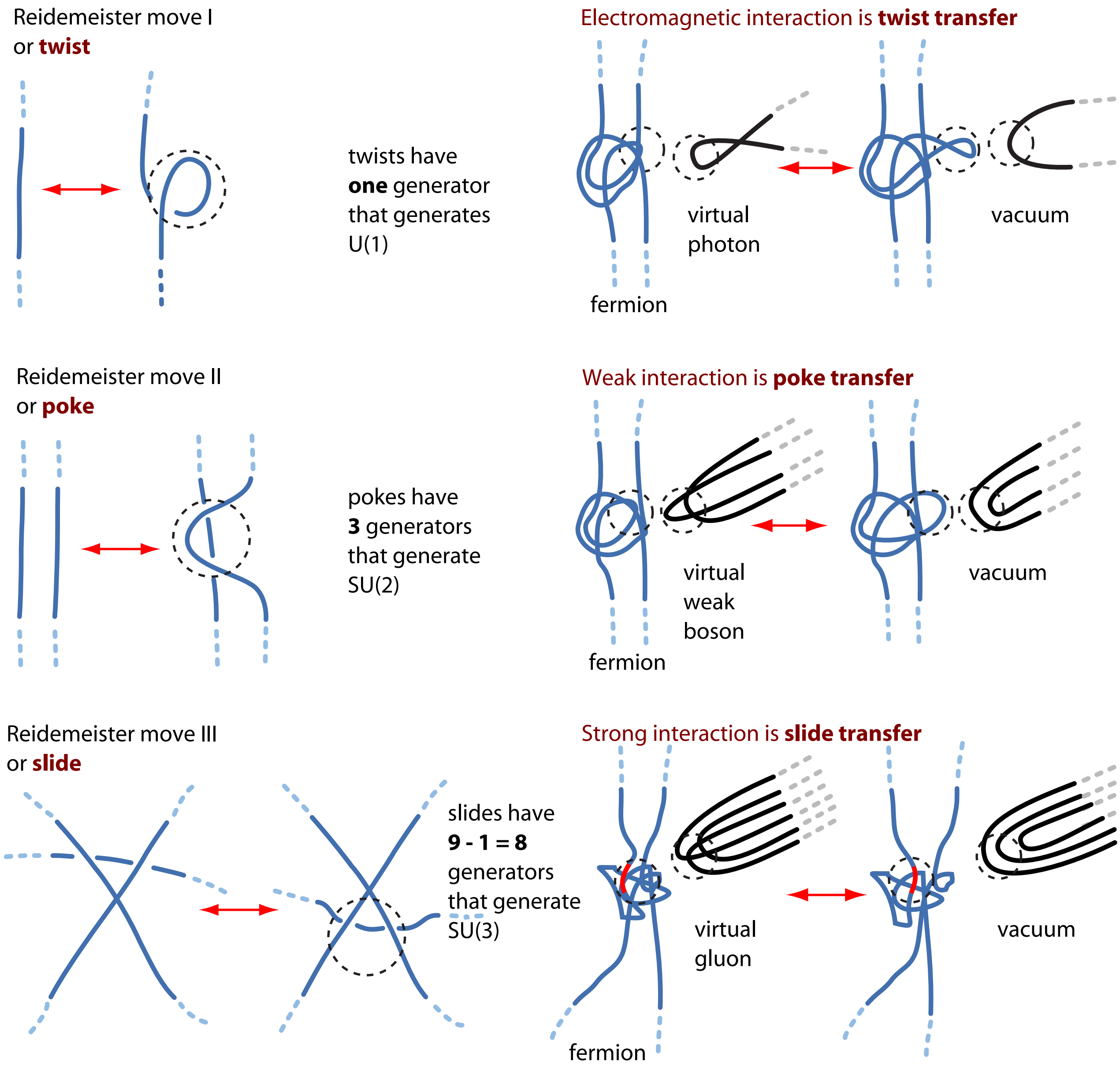


FIGURE 125 The three Reidemeister moves, i.e., the three possible deformations of tangle cores, determine the generators of the observed gauge interactions and thus determine their generator algebra. The generators rotate the regions enclosed by dotted circles by π. The full gauge groups arise by generalizing these rotations to continuous angles. Also the gauge coupling strengths arise.

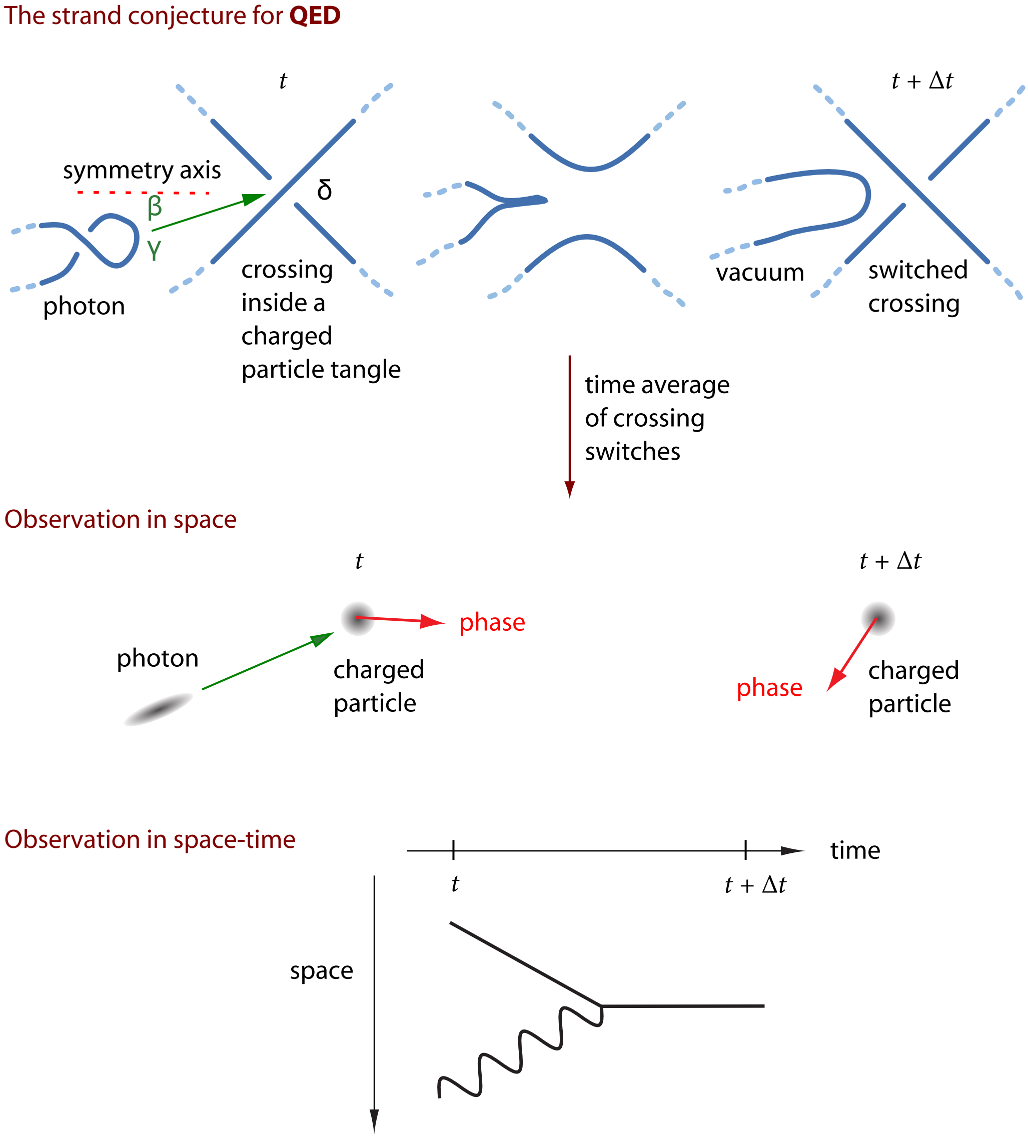


FIGURE 126 The geometric details of the basic process of quantum electrodynamics (QED). Top: the absorption of a photon by a tangle region carrying the charge $e/3$, viewed along the shortest distance of the crossing. Centre: the corresponding observation at usual experimental scales. Bottom: the basic Feynman diagram of QED. In short: twist transfer generates minimal coupling and determines the electromagnetic gauge coupling strength, thus the fine structure constant.

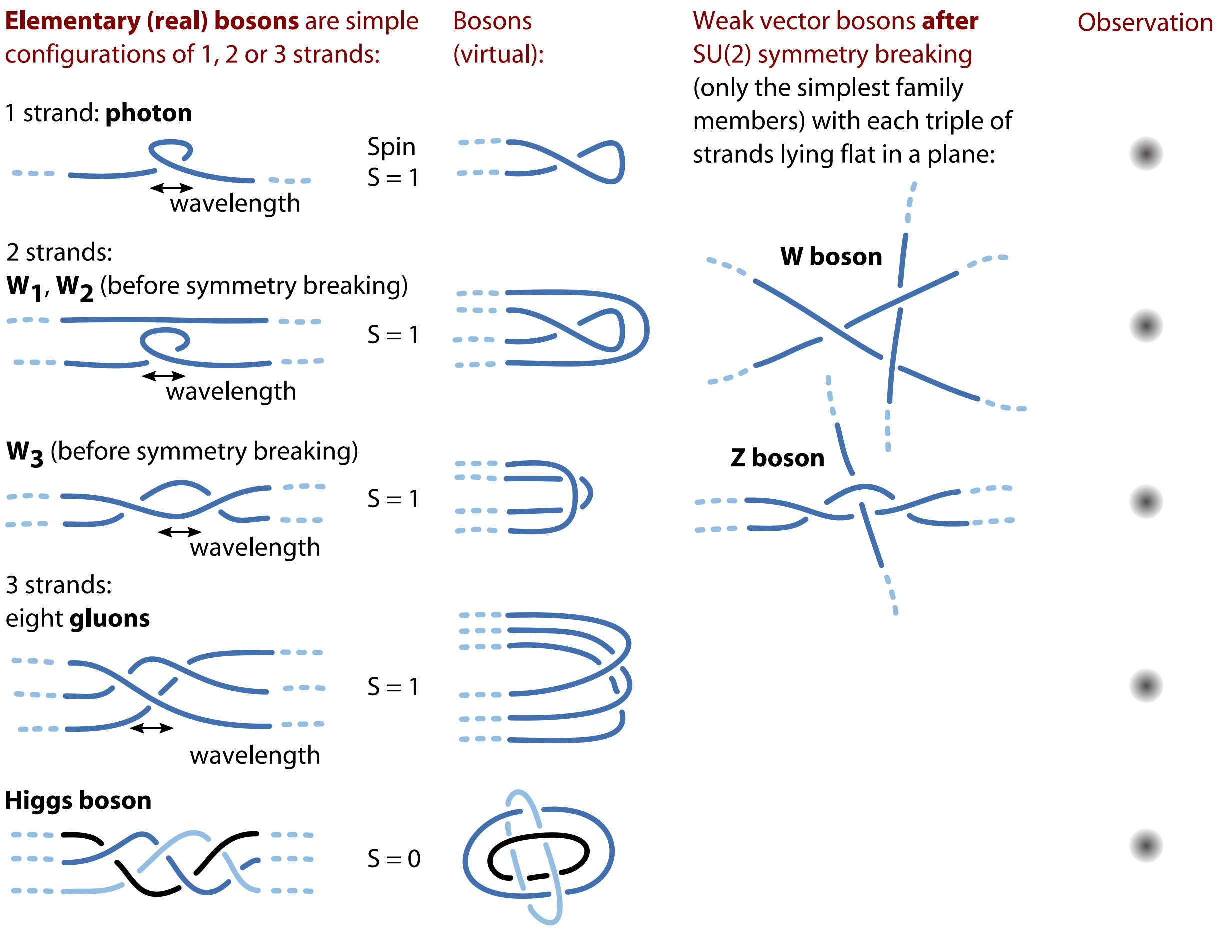


FIGURE 127 The tangle models for the elementary bosons. These tangles determine the spin values, the corresponding propagators, and ensure that the massless photons and gluons move with the speed of light. Apart from the graviton, no additional elementary boson appears to be possible. The tangle structure determines masses and the weak mixing angle.

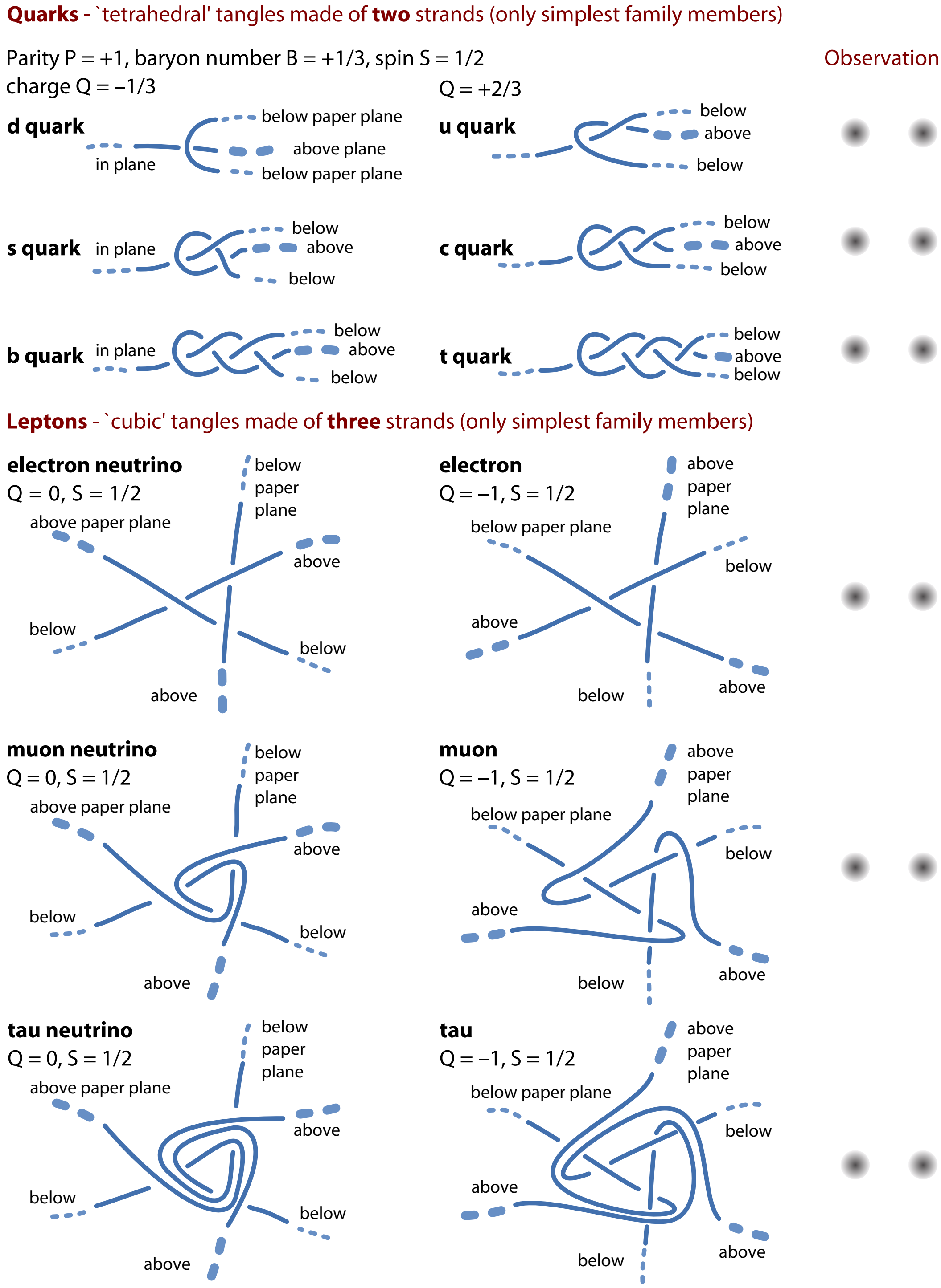


FIGURE 128 The *simplest* tangle models for the elementary fermions. Elementary fermions are described by rational, i.e., unknotted tangles. Their structures lead to coupling to the Higgs, as illustrated in Figure 130, produce positive mass values, and limit the number of generations to 3. Each fermion is represented by a tangle *family*, consisting of the simplest member and of tangles with added Higgs braids. The tangles determine the specific fermion propagators, including the values of masses and mixing angles. The tethers of the quark tangles follow the axes of a tetrahedron. The neutrino cores are simpler when seen in three dimensions: they are twisted triples of strands. The tethers of all lepton tangles approach the three coordinate axes at large distances from the core. No additional elementary fermions appear to be possible.

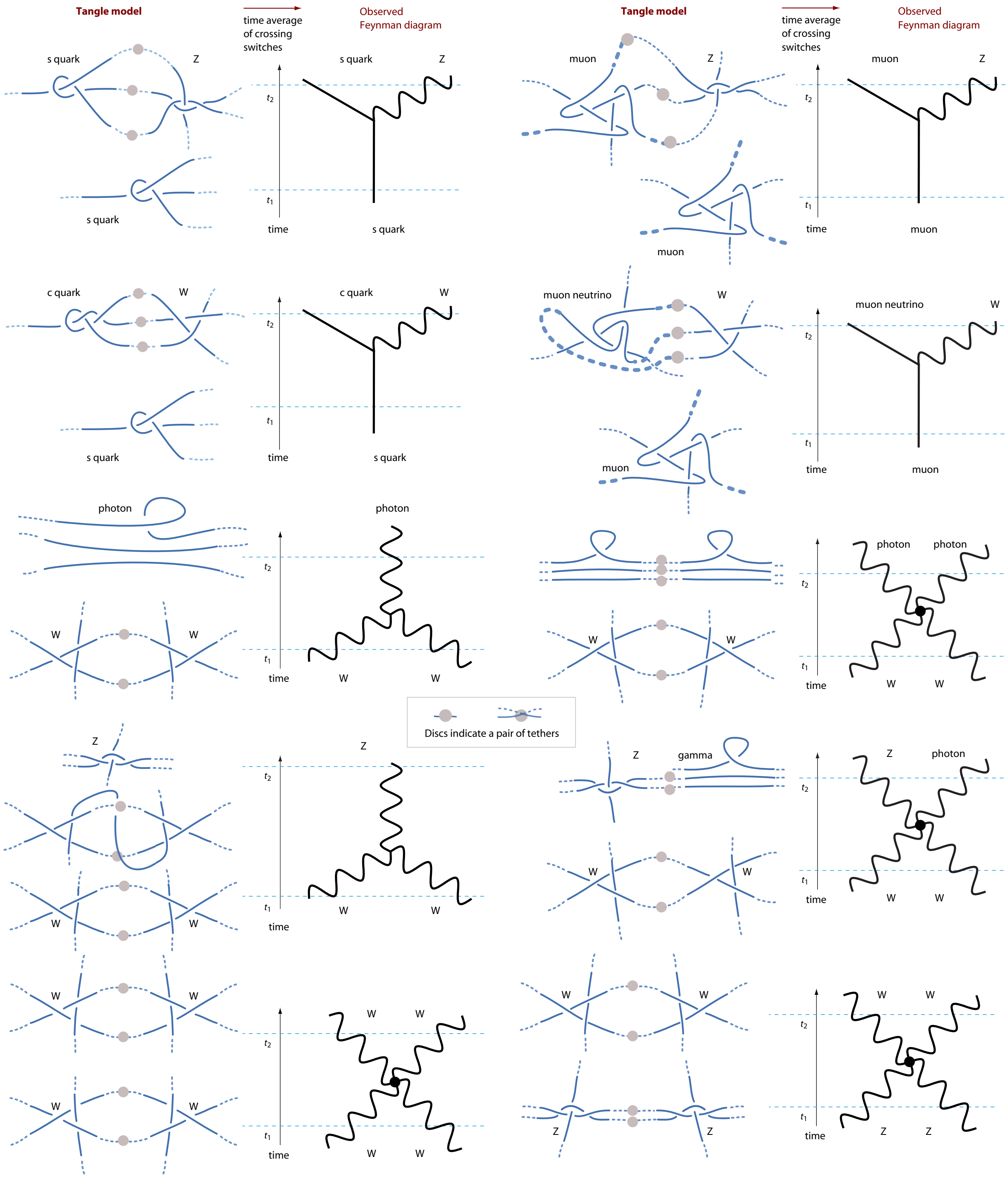


FIGURE 129 The vertices in Feynman diagrams allowed by the topology of the fermion and boson tangles (part one).

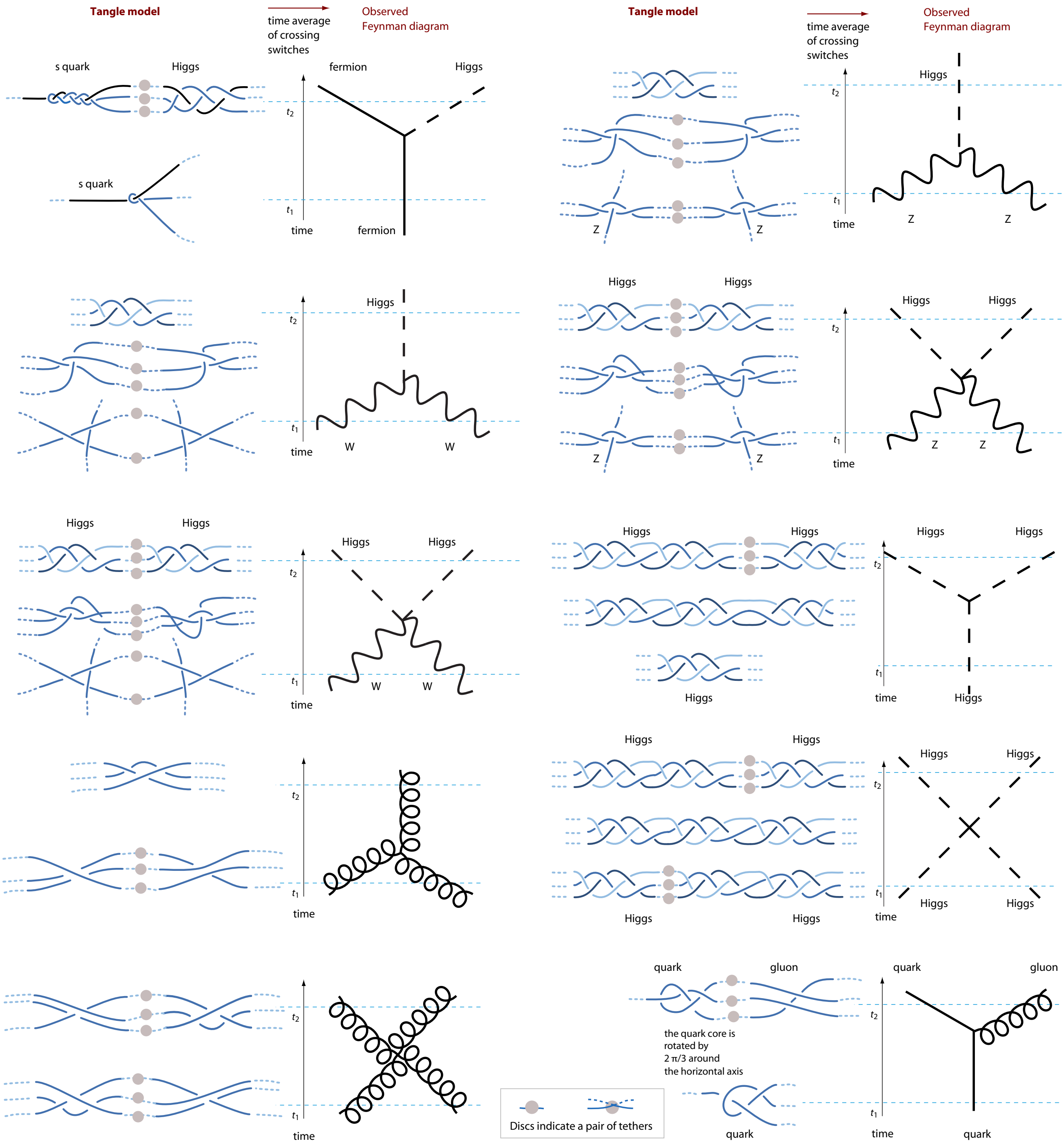
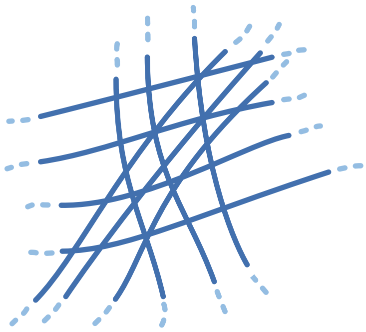


FIGURE 130 The vertices in Feynman diagrams allowed by the topology of the fermion and boson tangles (part two).

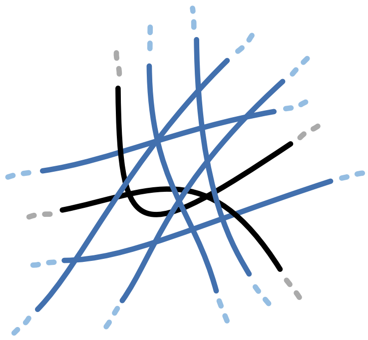
The strand conjecture for the **vacuum**



Observation	Predictions
Nothing (for long observation times)	Vanishing energy
Virtual pairs (for short observation times)	Emergent, Lorentz-invariant, and unique vacuum

FIGURE 131 An illustration of the strand conjecture for a flat vacuum: for sufficiently long time scales, the lack of crossing switches leads to a vanishing energy density; for short time scales, particle–antiparticle pairs, i.e., rational tangle–antitangle pairs, arise.

The strand conjecture for **curved space**



Observation	Predictions
Curved space	Black hole entropy
Non-trivial metric	Pure general relativity
Black holes	$P < c^5/4G$, $F < c^4/4G$
	Gravitons hard to detect

FIGURE 132 An illustration of the strand conjecture for a curved vacuum. The strand configuration is half way between that of a horizon and that of a flat vacuum. The black strands differ in their configuration from those in a flat vacuum.

The strand conjecture for the **graviton**

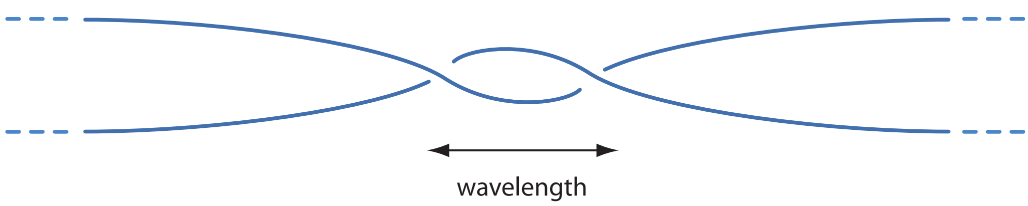


FIGURE 133 The strand conjecture for the graviton: a twisted pair of strands has spin 2, boson behaviour and vanishing mass.

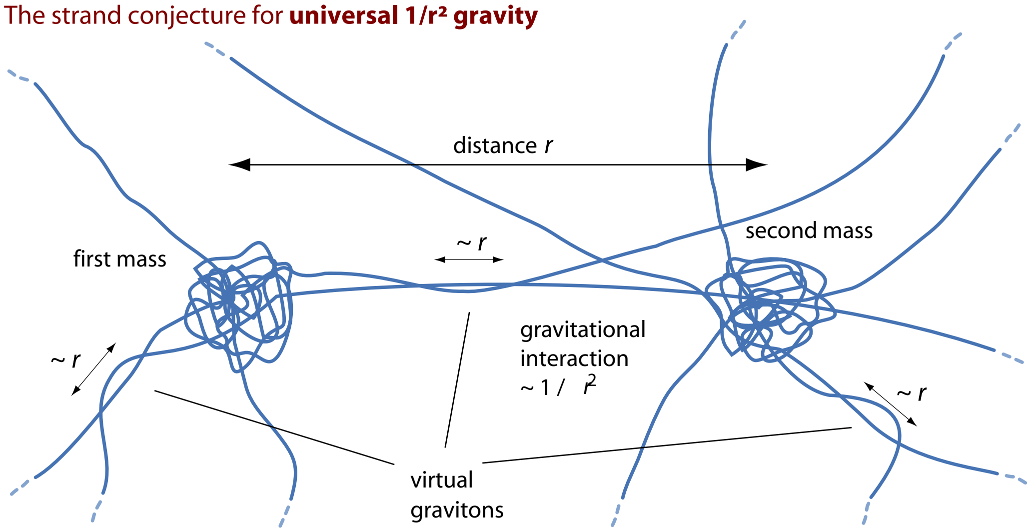


FIGURE 134 Gravitational attraction results from strands. For low speeds and negligible curvature, twisted tether pairs (virtual gravitons) from a mass lead to a $1/r^2$ attraction. The average length of twisted tether pairs scales with r and yields a $1/r^3$ decay of curvature. These conclusions are valid for infinite space without cosmological horizon.

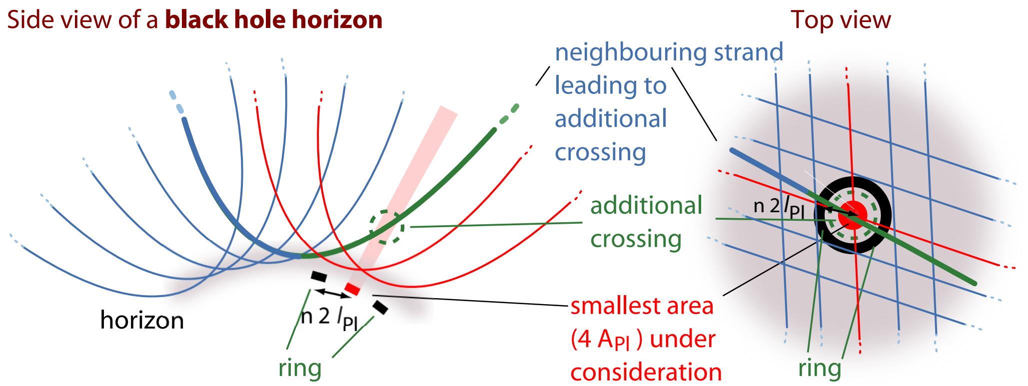


FIGURE 135 The strand conjecture for a Schwarzschild black hole: the black hole horizon is a cloudy or fuzzy surface produced by the crossing switches of the strands woven into it. Due to the additional crossings on the side of the observer, the number of micro-states per smallest area is larger than 2.

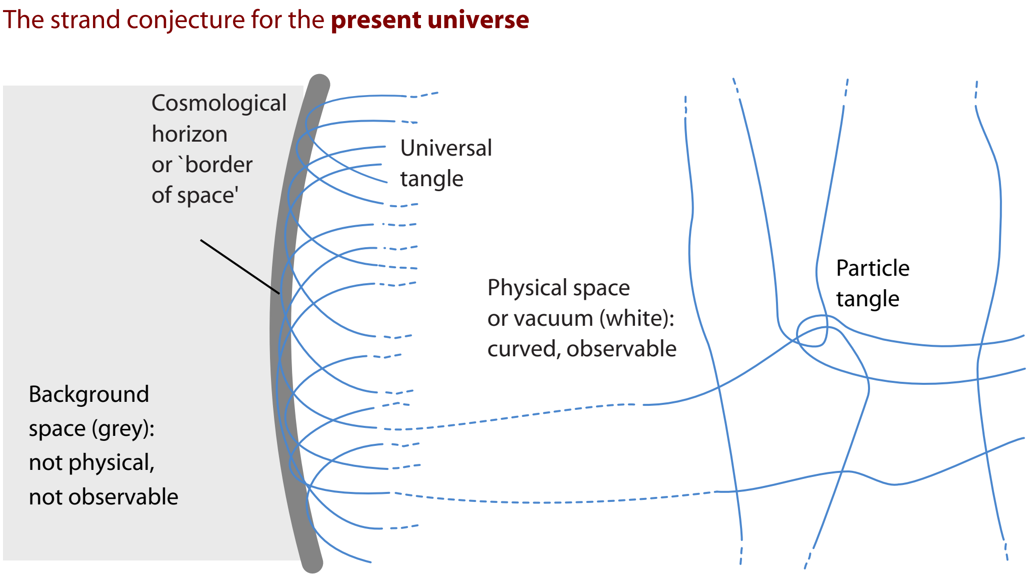


FIGURE 136 In the strand conjecture, the universe is limited by a cosmological (particle) horizon, as schematically illustrated here. Physical space (white) matches background space (grey) only inside the horizon. Physical space thus only exists *inside* the cosmic horizon.

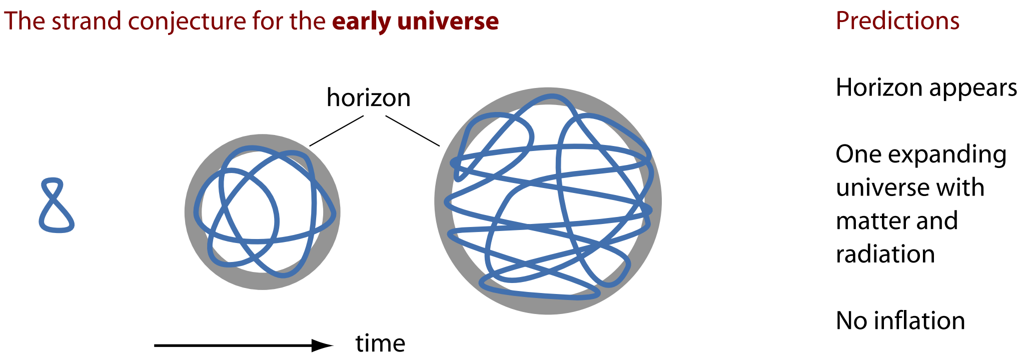


FIGURE 137 In the strand conjecture for the early universe, the universe increases in complexity over time and thereby forms a boundary: the cosmological horizon.

Chapter 14

EXPERIMENTAL PREDICTIONS OF THE STRAND MODEL

> Es gibt viele Theorien,
> die sich jedem Test entziehen.
> Diese aber kann man checken,
> elend wird sie dann verrecken.**
>
> Anonymous

Around the world, numerous researchers are involved in experiments that are searching for new effects. They are searching for new observations that are unexplained by the standard model of particle physics or by the conventional view of cosmology. At the same time, all these experiments are testing the strand model presented here. In fact, most people working on these experiments have not heard about the strand model, so that there is not even the danger of unconscious bias.

To simplify the check with experiments, the most important predictions of the strand model that we deduced in our adventure are listed in Table 17.

TABLE 17 The main predictions of the strand model that follow from the fundamental principle. The typeface distinguishes predictions that are unsurprising, that are *unconfirmed or unique* to the strand model, and those that are both **unconfirmed and unique**.

	Experiment	Prediction (most from 2008/2009)	Status (2019)
Page 37	Planck units (c, $\hbar$, k, $c^4/4G$)	are limit values.	None has been exceeded, but more checks are possible.
Page 330	Higgs boson	2012: does exist.	Verified.
Page 386	*Running of the coupling constants*	*2012: implies one Higgs.*	*Correct so far.*
Page 328 Page 330	*Longitudinal W and Z boson scattering*	*2012: show no non-local effects at the Large Hadron Collider.*	*None found yet.*
Page 328	*Longitudinal W and Z boson scattering*	*is unitary at the LHC.*	*Obvious.*
	W boson g-factor	is near to 2.	Is observed.
Page 352	*Unknown fermions (supersymmetric particles, magnetic monopoles, dyons, heavy neutrinos etc.)*	*do not exist.*	*None found yet.*

** No adequate translation is possible of this rhyme, inspired by Wilhelm Busch, claiming that any theory that can be tested is bound to die miserably.

TABLE 17 (Continued) The main predictions of the strand model that follow from the fundamental principle. The typeface distinguishes predictions that are unsurprising, that are *unconfirmed or unique* to the strand model, and those that are both **unconfirmed and unique**.

	Experiment	Prediction (most from 2008/2009)	Status (2019)
Page 314	*Unknown bosons (other gauge bosons, supersymmetric particles, axions etc.)*	*do not exist.*	*None found yet.*
Page 278, page 318	*Unknown interactions, energy scales and symmetries (grand unification, supersymmetry, quantum groups, technicolour etc.)*	*do not exist.*	*None found yet.*
Page 313	**Particle masses, mixing angles and coupling constants**	**are calculable by modifying existing software packages.**	**Most not yet calculated; approximations very encouraging.**
Page 313	Particle masses, mixing angles and coupling constants	are constant in time and space.	Is observed.
	Particle masses, mixing angles, coupling constants and g-factors	are identical for antimatter.	Is observed.
Page 376	Mixing matrix for quarks	is unitary.	Is observed.
Page 379	*Mixing matrix for neutrinos*	*is unitary.*	*No data yet.*
Page 339	*Neutrinos*	*are Dirac particles.*	*No data yet.*
Page 380	*Neutrinos*	*violate CP symmetry.*	*No data yet.*
Page 327	Neutrino-less double beta decay	does not exist.	Not yet found.
Page 278, page 327	**Electric dipole moments of elementary particles, magnetic dipole moment of neutrinos**	**have extremely small, calculable values.**	**No interesting data yet.**
Page 342	Tetraquarks	exist.	Likely.
Page 344, page 339	*Glueballs*	*probably do not exist; if they do, the spectrum can be compared to the strand model.*	*Not yet observed.*
Page 278, page 327	Proton decay and other rare decays, neutron-antineutron oscillations	occur at extremely small, standard model rates.	Not yet observed.
Page 340	Neutron decay	follows the standard model.	No deviations found.
Page 340	Neutron charge	vanishes.	None observed.
Page 336	*Hadron masses and form factors*	*can be calculated ab initio.*	*Not yet calculated; value sequences and signs correct.*
Page 355	Dark matter	is conventional matter plus black holes.	Partly confirmed by black hole mergers and lack of other results.
	Standard model of particle physics	*2012: is correct for all measurable energies.*	*All data agrees.*

TABLE 17 (Continued) The main predictions of the strand model that follow from the fundamental principle. The typeface distinguishes predictions that are unsurprising, that are *unconfirmed or unique* to the strand model, and those that are both **unconfirmed and unique**.

	Experiment	Prediction (most from 2008/2009)	Status (2019)
Page 148	Additional dimensions	do not exist.	Not observed.
Page 148	Non-commutative space-time	does not exist.	Not observed.
Page 295	General relativity	is correct at all accessible energies.	No deviation found.
Page 295	Short-distance deviations from universal gravitation and modified gravity	do not exist.	All data agrees.
Page 293	Space-time singularities, cosmic strings, wormholes, time-like loops, negative energy regions, domain walls	do not exist.	None observed.
Page 299, page 279	Quantum gravity effects	will not be found.	None observed yet.
Page 306	Behind a horizon	nothing exists.	Nothing observed.
Page 309	Cosmic inflation	did not occur.	Data not in contrast.
Page 380	Leptogenesis	did not occur.	Data are inconclusive.
Page 310	Cosmic topology	is trivial.	As observed.
Page 356	Vacuum	is stable and unique.	As observed.
	In summary: all motion	**results from strands.**	**Not yet falsified.**

In this list, the most interesting predictions of the strand model are the *numerical* predictions on the various mass ratios and mass sequences – including the Z/W and Higgs/W mass ratios – and the relative strength of the three gauge interactions. There is the clear option to calculate all fundamental constants in the foreseeable future.

In addition, the strand model reproduces the quark model, gauge theory, wave functions and general relativity; at the same time, the model predicts the lack of measurable deviations. The strand model solves conceptual problems such as the dark matter problem, inflation, confinement, the strong CP problem and the anomaly issue; by doing so, the strand model predicts the lack of unknown effects in these domains.

The strand model deduces all its experimental predictions from a single and simple fundamental principle: *events and Planck units are due to crossing switches of strands.* Provided there are no errors of reasoning, there is no way to change the predictions summarized here. The strand model is both simple and unmodifiable.

Naturally, errors of reasoning in the preceding chapters are well possible. A few have occurred in the past. The exploration was performed at high speed – possibly too high. If any experiment ever contradicts a prediction of the strand model, the model is doomed. When the above experimental predictions were first deduced in 2008 and 2009, they were quite unpopular. Practically all other attempts at unification predicted the existence of yet undiscovered particles and effects. However, so far, experiment does not confirm these other attempts; in fact, no prediction of the strand model has been falsified yet.

These predictions are not intellectual speculations. The author is prepared to take bets on each prediction of the above table – and has taken a few already.

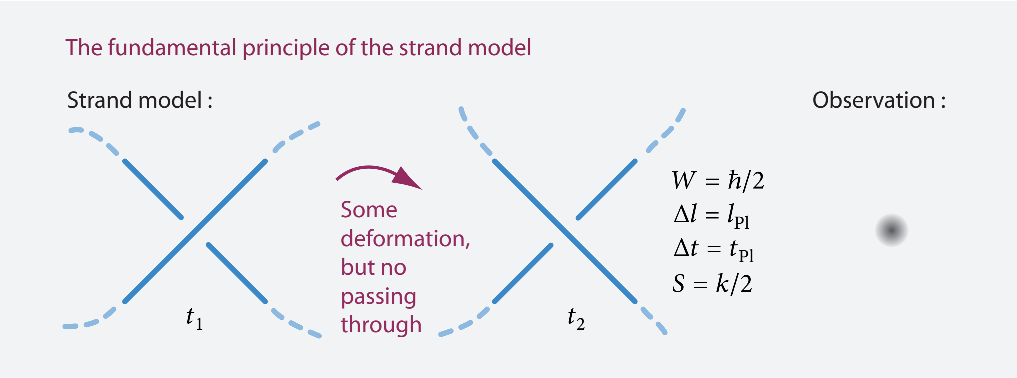


FIGURE 138 The fundamental principle of the strand model: *Planck units are defined by a crossing switch in three spatial dimensions.* With this principle, as shown in the previous chapters, the fundamental principle implies general relativity and the standard model of particle physics.

FINAL SUMMARY ABOUT THE MILLENNIUM ISSUES

In our adventure, we have argued that Planck's natural units should be modelled with the fundamental principle for strands, which is shown again in Figure 138. As we discovered, the fundamental principle explains the following measured properties of nature:

— Strands explain the principle of least action and the invariance of c, $\hbar$, G and k.
— Strands explain the three dimensions of space, the existence of gravitation, curvature and horizons, the equations of general relativity, the value of black hole entropy and the observations of modern cosmology.
— Strands explain all the concepts used in the Lagrangian of the standard model of particle physics, including wave functions, the Dirac equation and the finite, discrete and small mass of elementary particles.
— Strands explain the existence of electromagnetism and of the two nuclear interactions, with their gauge groups and all their other observed properties.
— Strands describe the observed gauge and Higgs bosons, their charges, their quantum numbers and their mass mass ranges.
— Strands explain the three generations of quarks and leptons, their charges and quantum numbers, their mixing, their mass sequences, as well as their confinement properties.
— Strands explain the quark model of hadrons, including CP violation, mass sequences, signs of quadrupole moments, the lack of unobserved hadrons, common Regge slopes and the existence of tetraquarks.
— Strands *do not allow* arbitrary values for masses, coupling constants, mixing angles and CP violating phases.
— Strands *enable* calculations of particle masses, their coupling constants, their mixing angles and the CP violating phases. First rough estimates of these values agree with the (much more precise) experimental data. Computer calculations will allow us to improve these checks in the near future.
— Strands *predict* the lack of unknown dark matter and of unknown inflation mechan-

isms.
— Finally, strands predict that nature *does not hide* any unknown elementary particle, fundamental interaction, fundamental symmetry or additional dimension. In particular, strands predict that no additional mathematical or physical concepts are required for a unified theory.

All these results translate to specific statements on experimental observations. So far, there is no contradiction between the strand model and experiments. These results allow us to sum up our adventure in three statements:

1. *Strands solve all open issues.* With one simple fundamental principle, the strand model solves or at least proposes a way to solve *all* issues from the millennium list of open issues in fundamental physics. All fundamental constants can be calculated. Page 19
2. *Strands agree with all observations.* In particular, the strand model implies that general relativity, quantum theory and the standard model of elementary particles are a *precise* description of motion for all practical purposes.
3. *Nothing new will be discovered in fundamental physics.* Unexpectedly but convincingly, strands predict that general relativity, quantum theory and the standard model of elementary particles are a *complete* description of motion for all practical purposes.

We have not yet literally reached the top of Motion Mountain – because certain numerical predictions of the fundamental constants are not yet precise enough – but if no cloud has played a trick on us, we have seen the top from nearby. In particular, we finally know the origin of colours.

The last leg, the accurate calculation of the constants of the standard model of particle physics, is still under way. The drive for simplicity and the spirit of playfulness that we invoked at the start have been good guides. Page 22

Chapter 15

THE TOP OF MOTION MOUNTAIN

> “All things are full of gods.
>
> Thales**”

Who am I? Where do I come from? What shall I do next? Where does the world come from? Can the whole world really come to a sudden end? What will happen in the future? What is beauty? All these questions have a common aspect: they are questions about motion. But what is motion? Our search for an answer led us to study motion in all its details. In this quest, every increase in the precision of our description of motion was a step towards the peak of Motion Mountain. Now that we arrived there, we can savour what we have achieved and recall the emotions that we have experienced.

In our ascent, we have learned how we move, how we experience our environment, how we grow, what parts we are made of, and how our actions and our convictions about them can be understood. We have learned a lot about the history and a bit about the future of matter, of radiation and of space. We have experienced and understood the many ways in which beauty appears in nature: as colours, as shapes, as rhythms and most of all: as simplicity.

Savouring our achievement means that first of all, we now can look back to where we came from. Then we enjoy the view we are offered and look out for what we could not see before. After that, we search for what is still hidden from our sight. And finally, we take a different path back down to where we live.

OUR PATH TO THE TOP

> “The labour we delight in physics pain.
>
> William Shakespeare, *Macbeth*.”

Our walk had a simple aim: to talk accurately about all motion. This 2500 year old quest drove us to the top of this mountain. We can summarize our path in three legs: everyday life, general relativity plus quantum theory, and unification.

** Thales of Miletus (*c.* 624 – *c.* 546 BCE) was the first known philosopher, mathematician and scientist.

Everyday life: the rule of infinity

Galilean physics is the description of everyday life. We all learned Galilean physics before secondary school. Galilean physics is the exploration and description of the motion of stones, water, trees, heat, the weather, electricity and light. To achieve this description of our environment, our first and main act in life is to partition experience into experience*s*. In other words, our first intellectual act is the invention of *parts*; we invented the *plural*.

Ref. 1, Ref. 3

The act of partitioning allows us to define sequences among our experiences, and thus to define the concept of *time*. The concept of *space* arises similarly by our possibility to distinguish observations that occur at the same time. By comparing parts with other parts, we define *measurement*. Using all of this, we become able to define *velocity*, *mass* and *electric charge*, among others. These allow us to introduce *action*, the quantity that quantifies change.

For a simple description of observations, we assume that division is possible without end: thus we introduce the infinitely small. We also assume that widening our scope of observation is possible without end. Thus we introduce the infinitely large. Defining parts thus leads us to introduce infinity.

Using parts and, with them, the infinitely small and the infinitely large, we found, in volumes I and III, that everyday motion has six main properties: it is continuous, conserved, relative, reversible, mirror-invariant and lazy. Motion is lazy – or efficient – because it produces as little change as possible.

Nature minimizes change. This is Galilean physics, the description of everyday motion, in one statement. It allows us to describe all our everyday experiences with stones, fluids, stars, electric current, heat and light. The idea of change-minimizing motion is based on a concept of motion that is continuous and predictable, and a concept of nature that contains the infinitely small and the infinitely large in every observable.

Relativity and quantum theory: the absence of infinity

> “Vorhin haben wir gesehen, daß in der Wirklichkeit das Unendliche nirgends zu finden ist, was für Erfahrungen und Beobachtungen und welcherlei Wissenschaft wir auch heranziehen.*”
>
> David Hilbert

The idea that nature offers an infinite range of possibilities is often voiced with deep personal conviction. However, the results of relativity and quantum theory show the opposite. In nature, speeds, forces, sizes, ages and actions are limited. No quantity in nature is infinitely large or infinitely small. No quantity in nature is defined with infinite precision. There never are infinitely many examples of a situation; the number of possibilities is always finite. The world around us is not infinite; neither its size, nor its age, nor its content. *Nature is not infinite.* This is general relativity and quantum theory in one statement.

Ref. 2, Ref. 4

Relativity and quantum theory show that the idea of infinity appears only in *approximate* descriptions of nature; it disappears when talking with precision. Nothing in nature

* ‘Above we have seen that in the real world, the infinite is nowhere to be found, whatever experiences and observations and whatever knowledge we appeal to.’

Ref. 263

is infinite. For example, we found in volume II that the sky is dark at night (also) because space is not infinite. And we found, in volumes IV and V, that quantum theory contains probabilities because there is a smallest action value in nature. In fact, the statement that a quantity is infinitely large or infinitely small cannot be confirmed or reproduced by any experiment. Worse, such a statement is falsified by every measurement. In short, we found that infinity is a fantasy of the human mind. In nature, it does not appear. *Infinity about nature is always a lie.*

The number of particles, their possible positions, the states they can have, our brain, our creativity, our possible thoughts: all this is not infinite. Nevertheless, quantum theory Ref. 4 and relativity changed the world: they allowed building ultrasound imaging, magnetic resonance imaging, lasers, satellite navigation systems, music players and the internet.

Despite the vast progress due to modern physics and the related technologies, one result remains: nothing in our environment is infinite – neither our life, nor our experiences, nor our memories, not even our dreams or our fantasies. Neither the information necessary to describe the universe, nor the paper to write down the formulae, nor the necessary ink, nor the time necessary to understand the formulae is infinite. Nature is not infinite. On the other hand, we also know that the illusion of the existence of infinity in nature is one the most persistent prejudices and myths ever conceived. Why did we use it in the first place?

The habit to use infinity to describe the world has many emotional reasons. For some, it reflects the deep-rooted experience of smallness that we carry within us as a remnant our personal history, when the world seemed so large and powerful. For others, the idea of our smallness allows us to deny somehow the responsibility for our actions or the existence of death. For others again, the idea of a finite universe often, at a first glance, produces deception, disbelief and discouragement. The absence of infinity means that we cannot achieve everything we want, and that our dreams and our possibilities are limited. Clinging to the idea of infinity is a way to avoid confronting this reality.

However, once we face and accept the absence of infinity, we make a powerful experience. We gain in strength. We are freed from the power of those who use this myth Challenge 221 e to put themselves above others. It is an illuminating experience to reread all those sentences on nature, on the world and on the universe containing the term 'infinite', knowing that they are incorrect, and then clearly experience the manipulations behind them. The desire to make others bow to what is called the infinite is a common type of human violence.

At first, the demise of infinity might also bring panic fear, because it can appear as a lack of guidance. But at closer inspection, the absence of infinity brings strength. Indeed, the elimination of infinity takes from people one of the deepest fears: the fear of being weak and insignificant.

Moreover, once we face the limits of nature, we react like in all those situations in which we encounter a boundary: the limit becomes a challenge. For example, the experience that all bodies unavoidably fall makes parachuting so thrilling. The recognition that our life is finite produces the fire to live it to the full. The knowledge of death gives meaning to our actions. In an infinite life, every act could be postponed without any consequence. The disappearance of infinity generates creativity. A world without limits is discouraging and depressing. Infinity is empty; limits are a source of strength and pour passion into our life. Only the limits of the world ensure that every additional step in life brings us forward. Only in a limited universe is progress possible and sensible. Who is wiser, the one who denies limits, or the one who accepts them? And who lives more intensely?

Unification: the absence of finitude

> Pray be always in motion. Early in the morning go and see things; and the rest of the day go and see people. If you stay but a week at a place, and that an insignificant one, see, however, all that is to be seen there; know as many people, and get into as many houses as ever you can.
>
> Philip Stanhope,* *Letters to his Son on the Fine Art of Becoming a Man of the World and a Gentleman.*

The last part of our adventure, described in this volume, produced an unexpected result. Not only is nature not infinite; nature is not finite either. None of the quantities which were supposed to be finite turn out to be so. Finitude turns out to be an approximation, or better, an illusion, though a subtle one. *Nature is not finite.* This is the unification of physics in one statement.

Precise observation shows that nothing in nature can be counted. If nature were finite it would have to be (described by) a set. However, the exploration of Planck scales shows that such a description is intrinsically incomplete and inaccurate. Indeed, a description of nature by a set can never explain the number of its elements, and thus cannot explain finitude itself. In other words, any approach that tries to describe nature as finite is a belief, and is never correct. *Finitude is a lie.*

Page 129

We thus lost our security of thought a second time. Nature is neither infinite nor finite. We explored the possibilities left over and found that only one option is left: *Nature is indivisible.* In other words, all parts that we experience are approximations. Both finitude and infinity are approximation of nature. All distinctions are approximate. This central conclusion solved the remaining open issues about motion. *Nature has no parts.*

The impossibility to count and the lack of parts imply that nature is not a computer, not an automaton, nor a physical system. *Nature is not discrete.*

Recognizing all distinctions as being approximate abolishes the distinction between the permanent aspects of nature ('objects', described by mass, charge, spin, etc.) and the changing aspects ('states', described by position, momentum, energy). Taking all distinctions as approximate introduces extended constituents: fluctuating strands. Looking even closer, these extended constituents are all the same one. Space, formally only used to describe states, also acquires changing aspects: it is made from fluctuating strands. Also properties like mass or charge, which formally were seen as static, become aspects of the ever changing interplay between these fundamental constituents. Describing nature as one fluctuating strand allows us to avoid finitude and to answer all questions left open by quantum theory and general relativity.

In a sense, the merging of objects and states is a resolution of the contrasting views on motion of the Greek thinkers Parmenides – 'there is no motion', i.e., in physical language, 'there are no states, there is only permanence' – and Heraclitus – 'everything

* Philip D. Stanhope (b. 1694 London, d. 1773 London) was a statesman and writer.

moves', i.e., in physical language 'there is no permanence, there are only states'. Both turn out to be right.

We can thus sum up the progress during our adventure of physics in the following table:

TABLE 18 The progress of physics.

Step 1	Galilean Physics	Nature is continuous.	We live in Galilean space.
Step 2	Relativity	Nature has no infinitely large.	We live in Riemannian space.
Step 3	Quantum field theory	Nature has no infinitely small.	We live in a Hilbert/Fock space.
Step 4	Unification	Nature is not finite. Nature has no parts.	We do not live in any space; we are space.

In summary, we are made of space. More precisely, we are made of the same constituents as space. In fact, the fascination of this result goes further than that.

NEW SIGHTS

> "Nel suo profondo vidi che s'interna,
> legato con amore in un volume,
> ciò che per l'universo si squaderna:
>
> sustanze e accidenti e lor costume
> quasi conflati insieme, per tal modo
> che ciò ch'i' dico è un semplice lume.
>
> La forma universal di questo nodo
> credo ch'i' vidi, perché più di largo,
> dicendo questo, mi sento ch'i' godo.*"
>
> Dante, *La Divina Commedia*, Paradiso, XXXIII, 85-93.

Modelling nature as a complicated web of fluctuating strands allowed us to describe at the same time empty space, matter, radiation, horizons, kefir, stars, children and all our other observations. All everyday experiences are consequence of everything in nature being made of one connected strand. This result literally widens our horizon.

* 'In its depth I saw gathered, bound with love into one volume, that which unfolds throughout the universe: substances and accidents and their relations almost joined together, in such a manner that what I say is only a simple image. The universal form of that knot, I think I saw, because, while I am telling about it, I feel deep joy.' This is, in nine lines, Dante's poetic description of his deepest mystical experience: the vision of god. For Dante, god, at the depth of the light it emanates, is a knot. That knot spreads throughout the universe, and substances and accidents – physicists would say: particles and states – are aspects of that knot. Dante Alighieri (b. 1265 Florence, d. 1321 Ravenna) was one of the founders and the most important poet of the Italian language. Most of the Divine Comedy, his magnum opus, was written in exile, after 1302, the year when he had been condemned to death in Florence.

The beauty of strands

> Someday, surely, we will see the principle underlying existence itself as so simple, so beautiful, so obvious, that we will all say to each other, "Oh, how could we all have been so blind, so long."
> John Wheeler, *A Journey Into Gravity And Spacetime.*

Describing everything as connected does not come natural to us humans. After all, in our life, we perform only one act: to partition. We define pluralities. There is no way we can avoid doing this. To observe, to think, to talk, to take a decision, to move, to suffer, to love or to enjoy life is impossible without partitioning.

Our walk showed us that there are limits to the ability to distinguish. Any kind of partitioning is always approximate. In fact, most people can summarize their personal experience by saying that they learned to make finer and finer distinctions. However, talking with highest precision about a part of the world inevitably leads to talk about the whole universe. The situation resembles a person who gets a piece of rope in his hand, and by following it, discovers a large net. He continues to pull and finally discovers that everything, including himself, is part of the net.

For the strand model, the term 'theory of everything' is therefore not acceptable. Nature cannot be divided into 'things'. In nature, things are never separable. There is no way to speak of 'every' thing; there are no sets, no elements and no parts in nature. A theory describing all of nature cannot be one of 'everything', as 'things' are only approximate entities: properly speaking, they do not exist. The strand model is not a theory of everything; it is a *complete theory*.

The strand model shows that nature is not made of related parts. Nature is made of relations only. Parts only exist approximately. The strand model also shows: being in motion is intrinsic to being a part. Parts, being approximate, are always in motion. As soon as we divide, we observe motion. The act of dividing, of partitioning, of defining parts is the very one which produces order out of chaos. Strands force us to rethink this habit.

Despite being so tough to grasp, strands yield a precise description of motion that unifies quantum field theory and general relativity. The strand model for the unification of motion is both simple and powerful. There are no free parameters. There are no questions left. Our view from the top of the mountain is thus complete. No uncertainty, no darkness, no fear and no insecurity are left over. Only wonder remains.

Can the strand model be generalized?

> Die Natur kann besser Physik als der beste Physiker.*
> Carl Ramsauer

Page 166 As mentioned above, mathematical physicists are fond of *generalizing* models. Despite this fondness, we required that any complete, unified description must be unique: any complete, unified description must be impossible to reduce, to modify or to generalize.

* 'Nature knows physics better than the best physicist.' Carl Ramsauer (b. 1879 Oldenburg, d. 1955 Berlin), influential physicist, discovered that electrons behave as waves.

In particular, a unified theory must neither be a generalization of particle physics nor of general relativity. Let us check this.

The strand model is not a generalization of general relativity: the definitions of curvature, of gravitons and of horizons differ radically from general relativity's approach. The strand model is also not a generalization of particle physics: the definitions of particle and of interactions differ radically from the concepts of quantum field theory. Indeed, we have shown that quantum field theory and general relativity are *approximations* to the strand model; they are neither special cases nor reductions of the strand model.

But what about the other requirements for a unified theory? Can the strand model be modified or generalized? We have seen that the model does not work in more spatial dimensions, does not work with more families of quarks, does not work with more interactions, and does not work with other evolution equations in general relativity or particle physics. The strand model does not work with other fundamental constituents, such as bifurcating entities, membranes, bands, or networks. (Though it does work with Page 172 the equivalent *funnels*, as explained earlier on, but that description is equivalent to the one with strands.) The strand model does not work with any modified fundamental principle. Obviously, exploring all possible variations and modifications remains a challenge Challenge 222 r for the years to come. If an actual modification of the strand model can be found, the strand model instantly loses its value: in that case, it would need to be shelved as a failure. Only a *unique* unified model can be correct.

In summary, one of the beautiful aspects of the strand model is its radical departure from twentieth-century physics in its basic concepts, combined with its almost incredible uniqueness. No generalization, no specialization and no modification of the strand Ref. 156 model seems possible. In short, the strand model qualifies as a unified, complete theory.

What is a requirement to one person, is a criticism to another. A number of researchers deeply dislike the strand model precisely because it doesn't generalize previous theories and because it cannot be generalized. This attitude deserves respect, as it is born from the admiration for several ancient masters of physics. However, the strand model points into a different direction.

What is nature?

> “Nature is what is whole in each of its parts.”
> Hermes Trismegistos, *Book of Twenty-four Philosophers*.

Ref. 264

At the end of our long adventure, we discovered that nature is not a set: everything is connected. Nature is only *approximately* a set. The universe has no topology, because space-time is not a manifold. Nevertheless, the approximate topology of the universe is that of an open Riemannian space. The universe has no definite particle number, because the universe is not a container; the universe is made of the same stuff of which particles are made. Nevertheless, the approximate particle density in the universe can be deduced.

In nature, everything is connected. This observation is reflected in the conjecture that all of nature is described by a single strand.

We thus arrive at the (slightly edited) summary given around the year 1200 by the author who wrote under the pen name Hermes Trismegistos: *Nature is what is whole in* *each of its parts*. But in contrast to that author, we now also know how to draw testable conclusions from the statement.

Quantum theory and the nature of matter and vacuum

> “In everything there is something of everything.”
> Anaxagoras of Clazimenes (500 –428 bce Lampsacus)

The strand model shows that as soon as we separate the universe into space-time and the rest, i.e., as soon as we introduce the coordinates x and t, quantum mechanics appears automatically. More precisely, *quantum effects are effects of extension*. Quantum theory appears when we realize that observations are composed of smallest events due to crossing switches, each with a change given by the quantum of action. All events and observations appear through the fluctuations of the strand that composes nature.

We found that *matter is made of tangled strands*. In fact, the correct way would be to say: matter is made of tangled strand *segments*. This connection leads to Schrödinger's equation and to Dirac's equation.

Insofar as matter is of the same fabric as the vacuum, we can rightly say that *everything is made of vacuum* and that *matter is made of nothing*. But the most appropriate definition arises when we realize that matter is not made from something, but that matter is a certain aspect of the *whole* of nature. Unification showed that every single elementary particle results from an arrangement of strands that involves the whole of nature, or, if we prefer, the entire universe. In other words, we can equally say: *matter is made of everything*.

We can also turn the equivalence of matter and vacuum around. Doing so, we arrive at the almost absurd statement: *vacuum is made of everything*.

> “Der heutigen Physik liegt die Frage nicht mehr ferne, ob nicht etwa alles, was ist, aus dem Äther geschaffen sei. Diese Dinge sind die äußersten Ziele unserer Wissenschaft, der Physik.*”
> Heinrich Hertz

Cosmology

The strand model also showed us how to deduce general relativity. The strand model clarified the fabric of horizons and explained the three dimensions of space. Most fascinating is the idea of a universe as the product of a single strand. A single strand implies that there was nothing before the big bang, and that there is nothing outside the night sky. For example, the strand model implies that there is no 'multiverse' and that there are no hidden worlds of any kind. And the fluctuating strand explains all observations of our universe.

The cosmological constant is not constant; it only measures the present age and size of
Page 8 the universe. Therefore, the constant does not need to appear in Figure 1. In other words, the cosmological constant simply measures the time from the big bang to the present.

* 'Modern physics is not far from the question whether everything that exists could possibly be made from aether. These things are the extreme goals of our science, physics.' Hertz said this in a well-known speech he gave in 1889. If we recall that 'aether' was the term of the time for 'vacuum', the citation is particularly striking.

The 'big bang' is the name for what we observe if we try to make observations approaching the limits of nature. The 'big bang' appears automatically from the strand model whenever we observe nature at the most distant times, the largest distances or at the largest energies: 'big bang' is the name for Planck scale physics.

The universe consists of a single strand. There are many particles in nature, because the strand is tangled up in complicated ways. What we call the 'horizon' of the universe is the place where new tangles appear.

The belief that the big bang or the horizon are examples of creation is incorrect. What happened at the big bang still happens at the horizon today. Both the black sky at night and the big bang are nature's way to tell us: 'Galilean physics is approximate! Quantum theory is approximate! General relativity is approximate!'

Musings about unification and strands

> "Continuing motion masters coldness.
> Continuing rest masters heat.
> Motion based on rest:
> Measure of the all-happening for the single one."
> Lao Tse,* *Tao Te King*, XXXXV.

All is made from one sort of thing: all is one substance. This idea, *monism*, sounds a lot like what the influential philosopher Baruch Spinoza (b. 1632 Amsterdam, d. 1677 The Hague) held as conviction. Monism, though mixed up with the idea of god, is also the basis of the philosophical ideas that Gottfried Wilhelm Leibniz (b. 1646 Leipzig, d. 1716 Hannover) presents in his text *La Monadologie*.

* *

Any complete theory of motion, also the strand model, is built on a single statement Ref. 265 about nature: The *many* exists only approximately. Nature is approximately multiple. The etymological meaning of the term 'multiple' is 'it has many folds'; in a very specific sense, nature thus has many folds.

* *

Any precise description of nature is free of arbitrary choices, because the divisions that we have to make in order to think are all common to everybody, and logically inescapable. Because physics is a consequence of this division, it is also 'theory-free' and 'interpretation-free'. This consequence of the complete theory will drive most philosophers up the wall.

* *

For over a century, physics students have been bombarded with the statement: 'Symmetries are beautiful.' Every expert on beauty, be it a painter, an architect, a sculptor, a musician, a photographer or a designer, fully and completely disagrees, and rightly so. Beauty has no relation to symmetry. Whoever says the contrary is blocking out his experiences of a beautiful landscape, of a beautiful human figure or of a beautiful work of art.

* Lao Tse (sixth century BCE) was an influential philosopher and sage.

The correct statement is: 'Symmetries simplify descriptions.' Symmetries simplify physical theories. That is the background for the statement of Werner Heisenberg: 'In Ref. 266 the beginning there was symmetry.' On the other hand, the strand model shows that even this statement is incorrect. In fact, neither the search for beauty nor the search for symmetry were the right paths to advance towards unification. Such statements have always been empty marketing phrases. In reality, the progress of fundamental theoretical physics was always driven by the search for *simplicity*.

* *

Strands unify physics. In particular, strands extend our views on quantum theory and mathematical physics, on particle physics and field theory, on axiomatic physics and algebraic physics, on polymer physics and gauge theory, on general relativity and cosmology. It will take several years before all these extensions will have been explored.

* *

The description of nature with strands is surprisingly simple, mainly because it uses so few basic concepts. Is this result astonishing? In our daily life, we describe our experiences with the help of a few thousand words, e.g. taking them from the roughly 350 000 words which make up the English language, or from a similar number from another language. This set is sufficient to talk about everything, from love to suffering, from beauty to happiness. And these terms are constructed from no more than about 35 basic Vol. III, page 280 concepts, as we have seen already. We should not be too surprised that we can in fact talk about the whole universe using only a few basic concepts: the act and the results of (approximate) distinction, or more specifically, a basic event – the crossing switch – and its observation.

* *

Almost all discoveries in physics were made at least 30 years too late. The same is true for the strand model. If we compare the strand model with what many physicists believed Page 22 in the twentieth century, we can see why: researchers had too many wrong ideas about unification. All these wrong ideas can be summarized in the following statement:

— 'Unification requires generalization of existing theories.'

This statement is subtle: it was rarely expressed explicitly but widely believed. But the statement is wrong, and it led many astray. On the other hand, the development of the strand model also followed a specific guiding idea, namely:

— 'Unification requires simplification.'

Hopefully this guiding idea will not become a dogma itself: in many domains of life, simplification means not to pay attention to the details. This attitude does a lot of harm.

* *

The strand model shows that achieving unification is not a feat requiring difficult abstraction. Unification was not hidden in some almost inaccessible place that can reached only by a few select, well-trained research scientists. No, unification is accessible to everyone who has a basic knowledge of nature and of physics. No Ph.D. in theoretical physics is

FIGURE 139 Motion Mountain does not resemble Cerro Torre, but a gentle hill (© Davide Brighenti, Myriam70)

needed to understand or to enjoy it. The knowledge presented in the previous volumes of this series is sufficient.

When Andrew Wiles first proved Fermat's last theorem after three centuries of attempts by the brightest and the best mathematicians, he explained that his search for a proof was like the exploration of a dark mansion. And seen the conceptual difficulties he had to overcome, the analogy was fitting. Recalling how many more people have already searched for unification without success, the first reaction is to compare the search for unification to the exploration of something even bigger, such as a complex dark cave system. But that analogy was not helpful. In contrast to the proof of Fermat's theorem, the goal of the quest for unification turned out to be simple and lying out in the open. Researchers had simply overlooked it, because they were convinced that the goal was complex, hidden in the dark and hard to reach. It was not.

The adventure of climbing Motion Mountain is thus not comparable to climbing Cerro Torre, which might be the toughest and most spectacular challenge that nature offers to mountain climbers. Figure 139 gives an impression of the peak. Motion Mountain does not resemble this peak at all. Neither does Motion Mountain resemble the Langtang Lirung peak in the Nepalese Himalayas shown on the cover of this volume. Climbing Motion Mountain is more like walking up a gentle green hill, alone, with a serene mind, on a sunny day, while enjoying the surrounding beauty of nature.

* *

Page 85 The strand model settles all questions about *determinism*. Quantum theory and general relativity are deterministic. Nevertheless, when both descriptions are combined, time turns out to be an approximate, low-energy concept. The same applies to determinism. Even though nature is deterministic for all practical purposes and shows no surprises, determinism shares the fate of all its conceivable opposites, such as fundamental randomness, indeterminism of all kinds, existence of wonders, creation out of nothing, or divine intervention: determinism is an *incorrect* description of nature at the Planck scale – like all its alternatives.

* *

The strand model also settles most so-called *really big questions* that John Wheeler used to ask: Why the quantum? How come existence? It from bit? A "participatory universe"? What makes "meaning"? Enjoy the exploration. Challenge 223 e

* *

Any unified model of nature encompasses a lot of ideas, issues and knowledge. Due to the sheer amount of material, publishing it in a journal will be challenging.

* *

Ref. 267 The strand model is so simple that it fits on a tombstone - or on a T-shirt. This would surely be god's favourite T-shirt. It is available at www.motionmountain.net/gfts.html.

* *

Historically, the strand model evolved from an exploration, started in the 1990s, of the maximum force in nature, the belt trick and the entropy of black holes. After the first six chapters of the present volume were completed in 2002, meditating on their implications led to the strand model and its fundamental principle.

Above all, it was the description of general relativity with the help of the maximum force that triggered the search for a unified description that was purely based on Planck units. Another essential point was the drive to search for a complete theory directly, from its requirements ('top down' in Figure 1), and *not* from the unification of quantum theory and general relativity ('bottom up'). Page 8 In the years from 2002 to 2007, most of the ideas of the strand model took shape, mainly in Munich's underground trains, while commuting between home and work. In those years, it appeared that strands could explain the Dirac equation, the entropy of black holes, general relativity and the particle spectrum with the three particle generations. While walking in the woods and fields around Munich, on 13 October 2008, it appeared that interactions are core deformations; in subsequent walks during 2008 and 2009, it appeared that strands explain the three gauge interactions, predict the lack of a Higgs boson – a bad mistake due to faulty reasoning, as turned out in 2012 – Page 330 and of any new physical effects beyond the standard model, and allow calculating the unexplained constants of particle physics. The model thus yielded almost all its main predictions before the accelerator experiments at the Large Hadron Collider at CERN in Geneva were switched on in autumn 2010. Thus much of the work was done in a haste – future will show what is of lasting value.

In 2012, the discovery of the Higgs boson, and in 2014, the comments by Sergei Fadeev led to an improvement and simplification of the strand model, eliminating knotted strands. From 2016 onwards, the experimental results of the LHC groups, of dark matter searches, and of the LIGO observatory confirmed the lack of deviations from the standard model of particle physics and from general relativity, as predicted by the strand model.

* *

Many researchers believed during all their life that the complete theory is something useful, important and valuable. This common belief about the importance and seriousness of the quest has led, over the past decades, to an increasingly aggressive atmosphere among these researchers. This unprofessional atmosphere, combined with the dependence of researchers on funding, has delayed the discovery of the complete theory by several decades.

In fact, the complete theory is *not useful:* it adds nothing of practical relevance to the combination of the standard model and general relativity. The unified theory is also *not important:* it has no application in everyday life or in industry and does not substantially change our view of the world; it just influences teaching – somewhat. Finally, the unified theory is *not valuable:* it does not help people in their life or make them happier. In short, the complete theory is what all fundamental theoretical research is: entertaining ideas.

Even if the strand model were to be replaced by another model, the conclusion remains: the unified theory is not useful, not important and not valuable. Knowing about unification does not confer any special powers. But it is enjoyable, comparable to a walk through a beautiful garden.

* *

The strand model will take a long time to get accepted. The first reason is obvious: *The strand model contradicts thinking habits* in many research fields. Researchers working on the foundations of quantum theory, on general relativity, on cosmic strings, on mathematical physics, on classical and quantum field theory, on polymer physics, on shape deformations, on quantum gravity, on strings, on the visualization of quantum mechanics, on knot theory, on higher dimensions, on supersymmetry, on the axiomatization of physics, on group theory, on the foundation of physics, on quantum optics and on particle physics have to give up many life-long thinking habits. So do all other physicists. *Strands supersede particles and points.*

There is also a second reason for the slow acceptance of the model presented here: *The strand model, in its simplicity, is only a small step away from present research.* Many researchers are finding out how close they have been to the ideas of the strand model, and for how long they were overlooking or ignoring such a simple option. The simplicity of the fundamental principle contrasts with the expectation of most researchers, namely that the unified theory is complicated, difficult and hard to discover. In fact, the opposite
Page 87
is true. *Strands are based on Planck units and provide a simple, almost algebraic description of nature.*

In summary, for many researchers and for many physicists, there is a mixture of confusion, anger and disappointment. It will take time before these feelings subside and are replaced by the clarity and fascination provided by the strand model.

“Only boring people get bored.”

Anonymous

The elimination of induction

> “Cum iam profeceris tantum, ut sit tibi etiam tui reverentia, licebit dimittas pedagogum.*”
>
> Seneca

The complete theory of motion has a consequence worth mentioning in detail: its lack of infinity and its lack of finitude eliminate the necessity of induction. This conclusion is of importance for general discussions on man's grasp of nature.

In physics, as in the other natural sciences, there is a tradition to state that a certain description of nature – once confusingly called a 'law' – is valid in *all* cases. In these statements, 'all' means 'for all values of the quantities appearing'. As a concrete example, the 'law' of universal gravitation is always claimed to be the same here and today, as well as at *all* other places and times, such as on the other end of the universe and in a few thousand years. The full list of such all-claims is part of the millennium list of open issues in twentieth-century physics. For many decades, the habit of claiming general validity from a limited and finite number of experiences, also called *induction*, has been seen, and rightly so, as a logically dubious manoeuvre, tolerated only because it works. But the developments described in this text show that this method is indeed justified.

Page 164

First of all, a claim of generality is not that enormous as it may seem, because the number of events that can be distinguished in nature is finite, not infinite. The preceding sections showed that the maximal number N of events that can be distinguished in the universe is of the order of $N = (T_0/t_{\mathrm{Pl}})^4 = 10^{244\pm2}$, T_0 being the age of the universe and t_{Pl} the Planck time. This is a big, but certainly finite number.

The unified description of nature has thus first reduced the various all-claims from an apparently infinite to a finite number of cases, though still involving astronomically large numbers. This reduction results from the recognition that infinities do not appear in the description of nature. We now know that when talking about nature, 'all' cases never means an infinite number.

A second, important result is achieved by the description of nature with strands. In any all-claim about fundamental motion, the checking of each of the large number of possibilities is not necessary any more, because all events result from a single entity, in which we introduce distinctions with our senses and our brain. And the distinctions we introduce imply automatically that the symmetries of nature – the 'all-claims' or 'inductions' – that are used in the description of motion are correct. Nature does not contain separate parts. Therefore, there is no way that separate parts can behave differently. Induction is a consequence of the unity of nature.

Ultimately, the possibility to *verify* statements of nature is due to the fact that all the aspects of our experience are *related*. Complete separation is impossible in nature. The verification of all-claims is possible because the strand model achieves the full description of how all 'parts' of nature are related.

The strand model shows that we can talk and think about nature because we are a part of it. The strand model also shows that induction works because everything in nature is related to everything else: nature is one.

* 'When you have profited so much that you respect yourself you may let go your tutor.' Seneca, the influential Roman poet and philosopher, writes this in his *Epistulae morales ad Lucilium*, XXV, 6.

What is still hidden?

> “That which eludes curiosity can be grasped in action.”
>
> Traditional saying.

Where do we come from? Where does the world come from? What will future bring? What is death? All these questions are questions about motion – and its meaning. To all such questions, the strand model does not provide answers. We are a collection of tangled strands. We are everything and nothing. The strand(s) we are made of will continue to fluctuate. Birth, life and death are aspects of tangled strands. The universe is a folded strand that grows in complexity.

Obviously, abstract statements about tangles do not help in any human quest. Indeed, we aimed at achieving a precise description of moving particles and bending space. Studying them was a sequence of riddles; but solving these riddles does not provide meaning, not even at the top of Motion Mountain. From the top we cannot see the evolution of complicated systems; in particular, we cannot see or describe the evolution of life, the biological evolution of species, or the growth of a human beings. Nor can we understand why we are climbing at all. Challenge 224 s

In short, from the top of Motion Mountain we cannot see the details down in the valleys of human relations or experience; strands do not provide advice or meaning. Remaining too long on the top is of no use. To find meaning, we have to descend back down to real life. Vol. I, page 16

A return path: je rêve, donc je suis

> “I hate reality. But it is the only place where one can get a good steak.”
>
> Woody Allen

Enjoying life and giving it meaning requires to descend from the top of Motion Mountain. The return path can take various different directions. From a mountain, the most beautiful and direct descent might be the use of a paraglider. After our adventure, we take an equally beautiful way: we leave reality.

The usual trail to study motion, also the one of this text, starts from our ability to talk about nature to somebody else. From this ability we deduced our description of nature, starting from Galilean physics and ending with the strand model. The same results can be found by requiring to be able to talk about nature to ourselves. Talking to oneself is an example of thinking. We should therefore be able to derive all physics from René Descartes' sentence ‘je pense, donc je suis’ – which he translated into Latin as ‘cogito ergo sum’. Descartes stressed that this is the only statement of which he is completely sure, in opposition to his observations, of which he is not. He had collected numerous examples in which the senses provide unreliable information. Ref. 268

However, when talking to ourselves, we can make more mistakes than when asking for checks from others. Let us approach this issue in a radically different way. We directly proceed to that situation in which the highest freedom is available and the largest number of mistakes are possible: the world of dreams. If nature would only be a dream, could we deduce from it the complete set of physical knowledge? Let us explore the issue.

— Dreaming implies the use of distinctions, of memory and of sight. Dreams contain *parts* and *motion*.
— Independently on whether dreams are due to previous observations or to fantasies, through memory we can define a sequence among them. The order relation is called *time*. The dream aspects being ordered are called *events*. The set of all (dream) events forms the (dream) *world*. Ref. 269
— In a dream we can have several independent experiences at the same time, e.g. about thirst and about hunger. Sequences thus do not provide a complete classification of experiences. We call the necessary additional distinction *space*. Dream space has three dimensions.* Dreaming thus means to use space and time.
— We can distinguish between dream contents. Distinguishing means that we can count items in dreams. Counting means that we have a way to define measurements. Dreams are thus characterized by something which we can call 'observables'. Dream experiences at a given instant of time are characterized by a *state*.
— Because we can describe dreams, the dream contents exist independently of dream time. We can also imagine the same dream contents at different places and different times in the dream space. There is thus an invariance of dream concepts in space and time. There are thus symmetries in dream space.
— Dream contents can interact. Dreams appear to vary without end. Dreams seem to be infinite.

In other words, a large part of the world of dreams is described by a modified form of *Galilean physics*. We note that the biggest difference between dreams and nature is the lack of conservation. In dreams, observations can appear, disappear, start and stop. We also note that instead of dreams, we could equally explore cinema *films*. Films, like dreams, are described by a modified form of Galilean physics. And films, like dreams, do not follow conservation laws. But dreams teach us much more.

— Dreams show that space can warp.
— Dream motion, as you may want to check, shows a maximum speed. Challenge 225 s
— Dreams show a strange limit in distance. There is a boundary to our field of vision, even though we do not manage to see it.

Pondering these issues shows that there are *limits* to dreams. In summary, the world of dreams has a maximum size, a maximum speed and three dimensions that can warp. The world of dreams and of films is described by a simple form of *general relativity*.

— Both the number of items we can dream of at the same time and the memory of previous dreams is finite.
— Dreams have colours.
— There are pixels in dreams, though we do not experience them directly. But we can do so indirectly: The existence of a highest number of things we can dream of at the same time implies that dream space has a smallest scale.

* Though a few mathematicians state that they can *think* in more than three spatial dimensions, all of them *dream* in three dimensions.

In summary, the world of dreams has something similar to a minimum change. The world of dreams and that of films is described by a simple form of *quantum theory*. The difference with nature is that in dreams and films, space is discrete from the outset. But there is still more to say about dreams.

— There is no way to say that dream images are made of mathematical points, as there is nothing smaller than pixels.
— In dreams, we cannot clearly distinguish objects ('matter') and environment ('space'); they often mix.
— In dreams, fluctuations appear both for images as well as for the background.
— In dreams, sharp distinctions are impossible. Dream space-time cannot be a set.
— Dream motion appears when approximate conservation (over time) is observed.
— In dreams, dimensionality at small distances is not clear; two and three dimensions are mixed up there.

In summary, the world of dreams seems to behave as if points and point particles do not exist; and since quantum theory and general relativity hold, the world of dreams seems to be described by extended constituents! We thus conclude this short exploration of the physics of dreams with a fascinating *conjecture*: even if nature would be a dream, an illusion or a fantasy, we might still get most of the results that we discovered in our ascent

Challenge 226 s of Motion Mountain. (What differences with modern physics would be left?) Speaking with tongue in cheek, the fear of our own faults of judgement, so rightly underlined by Descartes and many others after him, might not apply to fundamental physics.

WHAT IS THE ORIGIN OF COLOURS?

All colours around us are determined by the fine structure constant α – the coupling constant for the electromagnetic interaction at low energy – with its measured value

Ref. 5 1/137.035 999 139(31). The fine structure constant is also essential to describe most everyday devices and machines, as well as all human thoughts and movements. The constant is an aspect of every electric charge in nature.

The strand model showed us that electrical charge is a property of tangles of strands. In particular, the strand model showed:

▷ The fine structure constant describes the probability that a fluctuation adds a twist to the chiral tangles of electrically charged particles.

We have not yet deduced an accurate value for the fine structure constant, but we seem

Ref. 262 to have found out how to do so. In short, we seem to glimpse the origin of all colours – and thus of all beauty around us. Strands provide a beautiful explanation for beauty.

SUMMARY: WHAT IS MOTION?

> “Deep rest is motion in itself. Its motion rests in itself.”
> Lao Tse, *Tao Te King*, VI, as translated by Walter Jerven.

We can now answer the question that drove us through our adventure:

▷ ***Motion*** is the observation of crossing switches of the one, unobservable, tangled and fluctuating strand that describes all of nature.

Nature's strand forms particles, horizons and space-time: these are the parts of nature. Particles are tangles of strands; horizons and space-time are weaves of strands. The parts of nature move. The parts move because their strands fluctuate.

Motion appears because all parts in nature are approximate. Indeed, the observation of crossing switches and the description of strand segments fluctuating in a background space result and are possible because we approximate from the one strand that makes up nature to the many parts inside nature. The one strand (approximately) forms the many elementary particles inside us. Strand segments and particles (approximately) lead us to introduce background space, matter and radiation. Introducing background space implies observing motion. Motion thus appears automatically when approximate parts of nature, such as humans, animals or machines, describe other approximate parts of nature, such as other bodies or systems.

The observation of motion is due to our introduction of the plural. Motion results from of our forced use of *many (approximate) parts* to describe the *unity* of nature. The observation of motion results from approximations. All these approximate distinctions are unavoidable and are due to the limitations of our experience.

Motion appears as soon as we divide the world into parts and then follow these parts. Dividing nature into parts is not a conscious act; our human nature – our senses and our brain – *force* us to perform it. And whenever we experience or talk about parts of the universe, we find motion. Our senses and our brain are made to distinguish and to divide – and cannot do otherwise. We need to distinguish in order to survive, to think and to enjoy life. In a sense, we can say that motion appears as a logical consequence of our limitations; the fundamental limitation is the one that makes us distinguish and introduce parts, including points and sets.

Motion is an 'artefact' of locality. Locality is an approximation and is due to our human nature. Distinction, localization and motion are inextricably linked.

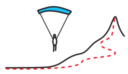
Motion is low energy concept. Motion does not exist at Planck scales, i.e., at the limits of nature.

Motion is an artefact due to our limitations. This conclusion resembles what Zeno of Elea stated 2500 years ago, that motion is an illusion. But in contrast to Zeno's pessimistic view, we now have a fascinating spectrum of results and tools at our disposition: they allow us to describe motion and nature with high precision. Most of all, these tools allow us to change ourselves and our environment for the better.

Vol. I, page 15

Ref. 270

> All the great things that have happened in the world first took place in a person's imagination, and how tomorrow's world will look like will largely depend on the power of imagination of those who are just learning to read right now.
>
> Astrid Lindgren*



* Astrid Lindgren (b. 1907 Näs, d. 2002 Stockholm) was a beloved writer of children books.

POSTFACE

Perhaps once you will read Plato's *Phaedrus*, one of the beautiful philosophical Greek texts. In it, Socrates is made to say that he almost never left the city walls because to him, as a 'lover of learning, trees and the open country do not teach anything, whereas men in the town do.' This is a veiled critique of Democritus, the most important and famous philosopher in Greece during Plato's time. Democritus was the natural philosopher par excellence, and arguably had learned from nature – with its trees and open country – more than anybody else after him.

After this adventure you can decide for yourself which of these two approaches is more congenial to you. It might be useful to know that Aristotle, Plato's pupil and the most influential Greek thinker, refused to choose and cultivated them both. There is no alternative in life to following one's own mind, and to enjoy doing so. If you enjoyed this particular trip, show it to your friends. For yourself, after this walk, sense intensively the pleasure of having accomplished something important. Many before you did not have the occasion. Enjoy the beauty of the view offered. Enjoy the vastness of horizon it provides. Enjoy the impressions that it creates inside you. Collect them and rest. You will have a treasure that will be useful in many occasions. Then, when you feel the desire of going further, get ready for another of the adventures life has to offer.

Plato's *Phaedrus*, written around 380 BCE, is available in many pocket editions. Do not waste your time learning ancient Greek to read it; the translated versions are as beautiful as the original.

Plato's lifelong avoidance of the natural sciences had two reasons. First of all, he was jealous of Democritus. Plato never even cites Democritus in his texts. Democritus was the most prolific, daring, admired and successful philosopher of his time (and maybe of all times). Democritus was a keen student of nature. His written works did not survive, because his studies were not congenial to the followers of christianity, and thus they were not copied by the monks in the Middle Ages. The loss of these texts is related to the second reason that kept Plato away from the natural sciences: he wanted to save his life. Plato had learned one thing from men in the town: talking about nature is dangerous. Starting around his lifetime, for over 2000 years people practising the natural sciences were regularly condemned to exile or to death for impiety. Fortunately, this is only rarely the case today. But such violence still occurs, and we can honour the dangers that those preceding us had to overcome in order to allow us enjoying this adventure.

Appendix A

KNOT AND TANGLE GEOMETRY

The following table provides a terse summary of the mathematics of knot shapes.

TABLE 19 Important properties of knot, links and tangles.

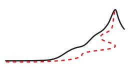
Concept	Defining property	Other properties
Knot / link / tangle	one closed / several closed / one or several open curves, all in 3d and without intersections	ropelength is integral of arclength; ropelength is shape-dependent.
Ideal knot, link, tangle (shape)	tightest possible knot, link or tangle (shape) assuming a rope of constant diameter that is infinitely flexible and infinitely slippery	at present, all non-trivial ideal shapes are only known approximately; most ideal knots (almost surely) have kinks.
Ribbon or *framing*	short perpendicular (or non-tangent) vector attached at each point of a curve	
Curvature of a curve	inverse curvature radius of 'touching' circle	measures departure from straightness, i.e., local bending of a curve.
Normal vector or curvature vector	local vector normal to the curve, in direction of the centre of the 'touching' circle, with length given by the curvature	is given by the second and first derivatives of the curve.
Binormal vector	local unit vector normal to the tangent and to the normal/curvature vector	
Torsion	local speed of rotation of the binormal vector; positive (negative) for right-handed (left-handed) helix	measures departure from flatness, i.e., local twisting or local handedness of a curve; essentially a third derivative of the curve.
Frenet frame at a curve point	'natural' local orthogonal frame of reference defined by *unit* tangent, *unit* normal/curvature and binormal vector	the Frenet frame differs at each curve point, the Frenet frame is *not* uniquely defined if the curve is locally straight.

TABLE 19 (Continued) Important properties of knot, links and tangles.

Concept	Defining property	Other properties
'Natural' framing or *Frenet ribbon*	defined by the local normal, i.e., local curvature vector	for a closed curve, it is always closed and two-sided, and thus never a Moebius band.
Linking number between two closed curves	sloppily, number of times that two curves wind around each other, or, equivalently, half the number of times that the curves 'swap' position	topological invariant, i.e., shape-independent; $\mathrm{Lk}(K1, K2) = \frac{1}{4\pi} \oint_{K2} \oint_{K1} \frac{\boldsymbol{r}_{12}\,(\mathrm{d}\boldsymbol{r}_1 \times \mathrm{d}\boldsymbol{r}_2)}{\boldsymbol{r}_{12}^3}$.
Linking number for a closed two-sided ribbon	number of times that the edges wind around each other	topological invariant, i.e., shape-independent; always an integer.
Self-linking number or 'natural' linking number for a knot	number of times that the edges of the natural/Frenet ribbon wind around each other	not a topological invariant, because of existence of inflection points.
Link integral for an open curve	generalization of the linking number for knots to open curves	usually not an integer.
Twist of a ribbon, open or closed	$\mathrm{Tw}(R)$ is the total angle, in units of 2π, by which the ribbon rotates around the central axis of the ribbon; sloppily said, it measures the *local* helicity; this type of twist has no relation to the first Reidemeister move	vanishes for ribbons that are everywhere flat.
Twist of a curve or knot	$\mathrm{Tw}(K)$ is the total angle, in units of 2π, by which the Frenet frame rotates around the tangent direction, or equivalently, (total) twist of the Frenet ribbon, also called the *total torsion* of the curve; this type of twist has no relation to the first Reidemeister move	not an integer even in case of knots; depends on curve/knot shape; is different from zero for chiral curves/knots; is zero for achiral curves/knots that have a rigid reflective symmetry; twist and torsion are only equal if the twist is defined with the Frenet ribbon – with other framings they differ.
Signed crossing number	sum of positive minus sum of negative crossings in a given oriented 2d projection of a curve or knot (sometimes called '2d-writhe')	always an integer; depends on shape.
2d-writhe of a knot, or *topological writhe*, or *Tait number*	signed crossing number for a *minimal* crossing number diagram/projection (sometimes the term '2d-writhe' is used for the signed crossing number of *any* configuration)	is shape-invariant; is always an integer; differs from 0 for all chiral knots; has the value 3 for the trefoil, 0 for the figure-eight knot, 5 for the 5_1 and 5_2 knots, 2 for the 6_1 knot, 7 for the 7_1 and 7_2 knots, 4 for the 8_1 knot, and 9 for the 9_2 knot.

TABLE 19 (Continued) Important properties of knot, links and tangles.

Concept	Defining property	Other properties
Writhing number or 3d-*writhe* of a knot	$\mathrm{Wr}(K)$ is the average, over all projection directions, of the signed crossing number; sloppily said, it measures how wrapped, coiled and chiral a knot is, i.e., it measures the *global* helicity	depends on knot shape; usually is not an integer; is different from zero for chiral knots; is zero for achiral knots that have a rigid reflective symmetry; $\mathrm{Wr}(K) = \frac{1}{4\pi} \oint_K \oint_K \frac{\boldsymbol{r}_{12}\,(\mathrm{d}\boldsymbol{r}_1 \times \mathrm{d}\boldsymbol{r}_2)}{r_{12}^3}$; uses no ribbon and thus is independent of the ribbon shape attached to the knot.
Writhe of ideal, alternating knots and of odd-component links	the value is quasi-quantized for alternating knots with small crossing numbers (< 11) in values that differ from $m4/7$ by only a few per cent	is additive under knot addition for knots with small crossing numbers (< 11) within less than 1 %.
Writhe of ideal, alternating even-component links	the value is quasi-quantized for alternating links with small crossing numbers (< 11) in values that differ from $2/7 + m4/7$ by only a few per cent	
Writhe of a ribbon	sloppily said, measures how wrapped, coiled and chiral a ribbon is, i.e., measures its *global* helicity	
Writhe of an open curve		vanishes for plane curves.
Calugareanu's theorem	for any knot K and any ribbon G attached to it, $\mathrm{Lk}(K, G) = \mathrm{Tw}(K, G) + \mathrm{Wr}(K)$	for applying the theorem to *open* curves, a (standardized) closing of curves is required.



CHALLENGE HINTS AND SOLUTIONS

Challenge 2, page 29: Take $\Delta f \ \Delta t \geqslant 1$ and substitute $\Delta l = c/\Delta f$ and $\Delta a = c/\Delta t$.

Challenge 16, page 44: Yes. But we can also argue its opposite, namely that matter appears when space is compressed too much. Both viewpoints are correct.

Challenge 22, page 47: The strictest upper limits are those with the smallest exponent for length, and the strictest lower limits are those with the largest exponent of length.

Challenge 24, page 49: To my knowledge, no such limits have been published. Do it yourself!

Challenge 25, page 49: The system limits cannot be chosen in other ways; after the limits have been corrected, the limits given here should still apply.

Challenge 28, page 50: Just insert numbers to check this.

Challenge 30, page 51: No.

Challenge 32, page 53: If you ever write such a table, publish it and send me a copy. I will include it in the text.

Challenge 35, page 66: Sloppily speaking, such a clock is not able to move its hands in a way that guarantees precise time reading.

Challenge 39, page 83: The final energy E produced by a proton accelerator increases with its radius R roughly as $E \sim R^{1.2}$; as an example, CERN's LHC achieved about 13 TeV for a radius of 4.3 km. Thus we would get a radius of more than 100 light years for a Planck energy accelerator. Building an accelerator achieving Planck energy is impossible.

Nature has no accelerator of this power, but gets near it. The maximum measured value of cosmic rays, 10^{22} eV, is about one millionth of the Planck energy. The mechanism of acceleration is still obscure. Neither black holes nor the cosmic horizon seem to be sources, for some yet unclear reasons. This issue is still a topic of research.

Challenge 40, page 84: The Planck energy is $E_{\rm Pl} = \sqrt{\hbar c^5/G} = 2.0$ GJ. Car fuel delivers about 43 MJ/kg. Thus the Planck energy corresponds to the energy of 47 kg of car fuel, about a tankful.

Challenge 41, page 84: Not really, as the mass error is equal to the mass only in the Planck case.

Challenge 42, page 84: It is improbable that such deviations can be found, as they are masked by the appearance of quantum gravity effects. However, if you do think that you have a prediction for a deviation, publish it, and send the author an email.

Page 281

Challenge 43, page 84: The minimum measurable distance is the same for single particles and systems of particles.

Challenge 44, page 85: There is no gravitation at those energies and there are no particles. There is thus no paradox.

Challenge 45, page 85: The issue is still being debated; a good candidate for a minimum momentum of a single particle is given by $\hbar/R$, where R is the radius of the universe. Is this answer satisfying?

Challenge 46, page 86: All mentioned options could be valid at the same time. The issue is not closed and clear thinking about it is not easy.

Challenge 47, page 86: The precise energy scale is not clear. The scale is either the Planck energy or within a few orders of magnitude from it; the lowest possible energy is thus around a thousandth of the Planck energy.

Challenge 49, page 88: If you can think of an experiment, publish the proposal, and send the author an email.

Vol. I, page 260 **Challenge 50**, page 91: The table of aggregates shows this clearly.

Challenge 51, page 92: The cosmic background radiation is a clock in the widest sense of the term.

Challenge 52, page 93: The way to deduce cosmological limits is presented in detail in the section starting on page 46.

Challenge 63, page 101: Also measurement errors at Planck scales prevent the determination of topology at those scales.

Challenge 65, page 103: The measurement error is as large as the measurement result.

Challenge 69, page 105: You will not find one.

Challenge 71, page 106: If you find one, publish it, and send the author an email.

Challenge 73, page 107: For the description of nature this is a contradiction. Nevertheless, the term 'universe', 'set of all sets' and other mathematical terms, as well as many religious concepts are of this type.

Challenge 74, page 109: No, for the reasons mentioned earlier on: fundamental measurement errors for horizon measurements, as well as many other effects, prevent this. The speculation is another example of misguided fantasy about extremal identity.

Challenge 75, page 109: The physical concepts most related to 'monad' are 'strand' and 'universe', as shown in the second half of this text.

Challenge 76, page 109: The macroscopic content of the universe may be observer-dependent. But to speak about many universes (Many 'everythings'?) or a 'multiverse' (What is more than everything? Why only one multiverse?) is pure nonsense. Vol. II, page 258

Challenge 79, page 109: True only if it were possible to do this. Because particles and space are indistinguishable, removing particles means to remove everything. (The strand model visualizes this connection most clearly.)

Vol. III, page 324 **Challenge 81**, page 109: True. Existence is the ability to interact. If the ability disappears, existence disappears. In other words, 'existence' is a low-energy concept.

Challenge 82, page 111: If you find a sensible statement about the universe, publish it! And send it to the author as well. The next challenge shows one reason why this issue is interesting. In addition, such a statement would contradict the conclusions on the combined effects of general relativity and quantum theory.

Challenge 83, page 111: Plotinus in the *Enneads* has defined 'god' in exactly this way. Later, Augustine in *De Trinitate* and in several other texts, and many subsequent theologians have taken up this view. (See also Thomas Aquinas, *Summa contra gentiles*, 1, 30.) The idea they propose is simple: it is possible to clearly say what 'god' is *not*, but it is impossible to say what 'god' *is*. This statement is also part of the official *Roman Catholic Catechism*: see part one, section one, chapter one, IV, 43, found at www.vatican.va/archive/ENG0015/__PC.HTM. Similar statements are found in Judaism, Hinduism and Buddhism.

In other terms, theologians admit that 'god' cannot be defined, that the term has no properties or content, and that therefore the term cannot be used in any positive sentence. The aspects

common to 'universe' and to 'god' suggest the conclusion that both are the same. Indeed, the analogy between the two concepts can be expanded to a proof: both concepts have the same content, the same boundary, and the same domain of application. (This is an intriguing and
Challenge 227 e fascinating exercise.) In fact, this might be the most interesting of all proofs of the existence of 'god', as it lacks all the problems that the more common 'proofs' have. Despite its interest, this proof of equivalence is not found in any book on the topic yet. The reason is twofold. First, the results of modern physics – showing that the concept of universe has all these strange properties – are not common knowledge yet. Secondly, the result of the proof, the identity of 'god' and 'universe' – also called *pantheism* – is a heresy for most religions. It is an irony that the catholic catechism, together with modern physics, can be used to show that pantheism is correct, because any catholic who defends pantheism (or other heresies following from modern physics) incurs automatic excommunication, latae sententiae, without any need for a formal procedure.

If one is ready to explore the identity of universe and 'god', one finds that a statement like 'god created the universe' translates as 'the universe implies the universe'. The original statement is thus not a lie any more, but is promoted to a tautology. Similar changes appear for many other – but not all – statements using the term 'god'. (The problems with the expression 'in the beginning' remain, though.) In fact, one can argue that statements about 'god' are only sensible and true if they remain sensible and true after the term has been exchanged with 'universe'.
Challenge 228 e Enjoy the exploration of such statements.

Challenge 85, page 114: If you find one, publish it and send it also to me. The conjecture is that no such effects exist.

Challenge 87, page 114: In fact, no length below the Planck length itself plays any role in nature.

Challenge 89, page 115: You need quantum humour, because the result obviously contradicts a previous one given on page 94 that includes general relativity.

Challenge 92, page 124: The number of spatial dimensions must be given first, in order to talk about spheres.

Challenge 93, page 128: This is a challenge to you to find out. It is fun, it may yield a result in contradiction with the arguments given so far (publish it in this case), or it may yield an independent check of the results of the section.

Challenge 95, page 132: This issue is open and still a subject of research. The conjecture of the author is that the answer is negative. If you find an alternative, publish it, and send the author an email.

Challenge 97, page 137: The lid of a box must obey the indeterminacy relation. It cannot be at perfect rest with respect to the rest of the box.

Challenge 99, page 138: No, because the cosmic background is not a Planck scale effect, but an effect of much lower energy.

Challenge 100, page 138: Yes, at Planck scales all interactions are strand deformations; therefore collisions and gravity are indistinguishable there.

Challenge 101, page 138: No. Time is continuous only if *either* quantum theory and point particles *or* general relativity and point masses are assumed. The argument shows that only the combination of *both* theories with continuity is impossible.

Challenge 102, page 138: You should, because at Planck scales nature's inherent measurement errors cannot clearly distinguish between different measurement results.

Challenge 103, page 138: We still have the chance to find the best approximate concepts possible. There is no reason to give up.

Challenge 104, page 138: Here are a few thoughts. A beginning of the big bang does not exist; something similar is given by that piece of continuous entity which is encountered when going backwards in time as much as possible. This has several implications.

— Going backwards in time as far as possible – towards the 'beginning' of time – is the same as zooming to smallest distances: we find a single strand of the amoeba.
— In other words, we speculate that the whole world is one single piece, fluctuating, and possibly tangled, knotted or branched.
— Going far away into space – to the border of the universe – is like taking a snapshot with a short shutter time: strands everywhere.
— Whenever we sloppily say that extended entities are 'infinite' in size, we only mean that they reach the horizon of the universe.

In summary, no starting point of the big bang exists, because time does not exist there. For the same reason, no initial conditions for particles or space-time exist. In addition, this shows that the big bang involved no creation, because without time and without possibility of choice, the term 'creation' makes no sense.

Challenge 105, page 138: The equivalence follows from the fact that all these processes require Planck energy, Planck measurement precision, Planck curvature, and Planck shutter time.

Page 369 **Challenge 106**, page 138: No, as explained later on in the text.

Challenge 107, page 139: Probably there is nothing wrong with the argument. For example, in the strand model, all observables are composed of fundamental events, and so, in some way, all observables are fundamentally indistinguishable.

Challenge 108, page 139: If not, force yourself. Brainstorming is important in life, as is the subsequent step: the checking of the speculations.

Challenge 113, page 151: The author would like to receive a mail on your reasons for disagreement.

Challenge 114, page 153: Let the author know if you succeed. And publish the results.

Challenge 115, page 153: Energy is action per time. Now, the Planck constant is the unit of action, and is defined by a crossing switch. A system that continuously produces a crossing switch for every Planck time running by thus has Planck energy. An example would be a tangle that is rotating extremely rapidly, once per Planck time, producing a crossing switch for every turn.

Momentum is action per length. A system that continuously produces a crossing switch whenever it advances by a Planck length has Planck momentum. An example would be a tangle configuration that lets a switch hop from one strand to the next under tight strand packing.

Force is action per length and time. A system that continuously produces a crossing switch for every Planck time that passes by and for every Planck length it advances through exerts a Planck force. A tangle with the structure of a screw that rotates and advances with sufficient speed would be an example.

Challenge 119, page 163: Yes; the appearance of a crossing does not depend on distance or on the number of strands in between.

Challenge 120, page 163: No; more than three dimensions do not allow us to define a crossing switch.

Challenge 121, page 163: If so, let the author know. If the generalization is genuine, the strand model is not correct.

Challenge 133, page 190: The magnitude at a point should be related to the vectorial sum of all inverse shortest crossing distances at that point.

Challenge 139, page 201: This algebraic transformation is shown in all textbooks that treat the Pauli equation. It can also be checked by writing the two equations out component by component.

Challenge 143, page 222: Yes, as can easily be checked by rereading the definitions with the spinor tangle description in mind.

Challenge 146, page 222: No contradiction is known.

Challenge 147, page 222: In the relativistic case, local space curvature is also taken into account.

Challenge 149, page 222: Find out, publish the result, and let the author know.

Challenge 150, page 223: If the strand interpenetration is allowed *generally*, quantum theory is impossible to derive, as the spinor behaviour would not be possible. If strand interpenetration were allowed only *under certain conditions* (such as only for a strand with itself, but not among two different strands), quantum theory might still possible. A similar process lies at the basis of mass generation, as shown in the section on the weak interaction.

Challenge 152, page 223: The belt trick would imply that a wheel rolls over its own blood supply at every second rotation.

Challenge 161, page 245: The author bets that you cannot find a deviation of the strand model from QED. If you find one, publish it!

Challenge 172, page 274: No slide is possible, thus no crossing change appears; thus the situation has no observable effects. If we deform one slide before the slide – which is possible – we get back the situation already discussed above.

Challenge 177, page 275: For the Wightman axioms, this seems to be the case; however, a formal proof is still missing. The same is expected for the Haag–Kastler axioms.

Challenge 187, page 303: A black hole has at least one crossing, thus at least a Planck mass.

Challenge 193, page 324: These tangles are not rational. In the renewed strand model of 2015, they cannot form; they are not allowed and do not represent any particle.

Challenge 195, page 339: Such a tangle is composed of several gravitons.

Challenge 196, page 346: Tail braiding leads to tangledness, which in turn is the basis for core rotation. And core rotation is kinetic energy, not rest mass.

Challenge 198, page 346: The issue is topic of research; for symmetry reasons it seems that a state in which each of the six quarks has the same bound to the other five quarks cannot exist.

Challenge 205, page 372: If you find such an estimate, publish it and send it to the author. A really good estimate also answers the following question: why does particle mass increase with core complexity? A tangle with a complex core, i.e., with a core of large ropelength, has a large mass value. Any correct estimate of the mass must yield this property. But a more complex knot will have a smaller probability for the belt trick. We seem to be forced to conclude that particle mass is not due to the belt trick alone.

Challenge 206, page 373: If you find such an estimate, publish it and send it to the author.

Challenge 208, page 373: Probably not.

Challenge 209, page 373: Probably not.

Challenge 210, page 373: Probably not.

Challenge 211, page 374: Find out – and let the author know.

Challenge 213, page 378: This would be an interesting result worth a publication.

Challenge 216, page 386: If you plan such a calculation, the author would be delighted to help.

Challenge 220, page 395: Take up the challenge!

Challenge 222, page 423: There is a good chance, however, that such alternatives can be eliminated rather quickly. If you cannot do so, do publish the argument, and let the author know.

Challenge 224, page 431: Nobody can really answer 'why'-questions about human actions. Climbing, like every other passion, is also a symbolic activity. Climbing can be a search for adventure, for meaning, for our mother or father, for ourselves, for happiness, or for peace.

Challenge 225, page 432: Also in dreams, speeds can be compared; and also in dreams, a kind of causality holds (though not a trivial one). Thus there is an invariant and therefore a maximum speed.

Challenge 226, page 433: Probably none. The answer depends on whether the existence of strands can be deduced from dreams. If strands can be deduced from dreams, all of physics follows. The conjecture is that this deduction is possible. If you find an argument against or in favour of this conjecture, let the author know.

BIBLIOGRAPHY

> The only end of writing is to enable the readers better to enjoy life, or better to endure it.
>
> Samuel Johnson*

1 See the first volume of the Motion Mountain series, *Fall, Flow and Heat*, available as free download at www.motionmountain.net. Cited on pages 17 and 418.

2 See the second volume of the Motion Mountain series, *Relativity*, available as free download at www.motionmountain.net. Cited on pages 17, 18, 418, and 448.

3 See the third volume of the Motion Mountain series, *Light, Charges and Brains*, available as free download at www.motionmountain.net, as well as the mentioned fourth and fifth volumes. Cited on pages 17 and 418.

4 See the fourth and fifth volumes of the Motion Mountain series, *The Quantum of Change* and *Pleasure, Technology and the Stars*, available as free download at www.motionmountain.net. Cited on pages 18, 418, 419, and 447.

5 The most precise value of the fine structure constant is determined from a weighted world average of high-precision measurements by a special international scientific committee called CODATA. Its website is www.codata.org/committees-and-groups/fundamental-physical-constants. The site also provides the latest official publication with the values of the fundamental constants. The most recent value of the fine structure constant is published at physics.nist.gov/cgi-bin/cuu/Value?alphinv and physics.nist.gov/cgi-bin/cuu/Value?alph. Cited on pages 18, 229, 382, 391, and 433.

6 See for example, the book by ROBERT LAUGHLIN, *A Different Universe: Reinventing Physics from the Botton Down* Basic Books, 2005. Of the numerous books that discuss the idea of a final theory, this is the only one worth reading, and the only one cited in this bibliography. The opinions of Laughlin are worth pondering. Cited on page 21.

7 Many physicists, including Steven Weinberg, regularly – and incorrectly – claim in interviews that the measurement problem is not solved yet. Cited on page 21.

8 Undocumented sentences to this effect are regularly attributed to Albert Einstein. Because Einstein was a pantheist, as he often explained, his statements on the 'mind of god' are not really to be taken seriously. They were all made – if at all – in a humorous tone. Cited on page 21.

9 For an example for the inappropriate fear of unification, see the theatre play *Die Physiker* by the Swiss author FRIEDRICH DÜRRENMATT. Several other plays and novels took over this type of disinformation. Cited on page 22.

* This is a statement from the brilliant essay by the influential writer SAMUEL JOHNSON, *Review of Soame Jenyns' "A Free Enquiry Into the Nature and Origin of Evil"*, 1757. See www.samueljohnson.com.

10 Exploring the spirit of play is the subject of research of the famous National Institute for Play, founded by Stuart Brown, and found at www.nifplay.org. Cited on page 22.

11 See e.g. the 1922 lectures by Lorentz at Caltech, published as H. A. Lorentz, *Problems of Modern Physics*, edited by H. Bateman, Ginn and Company, 1927, page 99. Cited on page 28.

12 Bohr explained the indivisibilty of the quantum of action in his famous Como lecture, printed in N. Bohr, *Atomtheorie und Naturbeschreibung*, Springer, 1931. It was translated into English language as N. Bohr, *Atomic Theory and the Description of Nature*, Cambridge University Press, 1934. More statements about the indivisibility of the quantum of action can be found in N. Bohr, *Atomic Physics and Human Knowledge*, Science Editions, New York, 1961. For summaries of Bohr's ideas by others see Max Jammer, *The Philosophy of Quantum Mechanics*, Wiley, first edition, 1974, pp. 90–91, and John Honner, *The Description of Nature – Niels Bohr and the Philosophy of Quantum Physics*, Clarendon Press, 1987, p. 104. Cited on page 29.

13 For an overview of the quantum of action as a basis of quantum theory, see the first chapter of the fourth volume of the Motion Mountain series, Ref. 4. Cited on page 30.

Vol. IV, page 15

Vol. IV, page 182

14 An overview of EBK quantization can be found in the volume on quantum theory. Cited on page 30.

15 Minimal entropy is discussed by L. Szilard, *Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen*, Zeitschrift für Physik **53**, pp. 840–856, 1929. This classic paper can also be found in English translation in his collected works. Cited on page 31.

16 See for example A. E. Shalyt-Margolin & A. Ya. Tregubovich, *Generalized uncertainty relation in thermodynamics*, preprint at arxiv.org/abs/gr-qc/0307018, or J. Uffink & J. van Lith-van Dis, *Thermodynamic uncertainty relations*, Foundations of Physics **29**, pp. 655–692, 1999. Cited on page 31.

17 See also the fundamental paper by A. DiSessa, *Momentum flow as an alternative perspective in elementary mechanics*, **48**, p. 365, 1980, and A. DiSessa, *Erratum: "Momentum flow as an alternative perspective in elementary mechanics" [Am. J. Phys. 48, 365 (1980)]*, **48**, p. 784, 1980. Cited on page 32.

18 The observations of black holes at the centre of galaxies and elsewhere are summarised by R. Blandford & N. Gehrels, *Revisiting the black hole*, Physics Today **52**, June 1999. Their existence is now well established. Cited on page 32.

19 It seems that the first published statement of the maximum force as a *fundamental principle* was around the year 2000, in this text, in the chapter on gravitation and relativity. The author discovered the maximum force principle, not knowing the work of others, when searching for a way to derive the results of the last part of this adventure that would be so simple that it would convince even a secondary-school student. In the year 2000, the author told his friends in Berlin about his didactic approach for general relativity.

Page 57

The *concept* of a maximum force was first proposed, most probably, by H.-J. Treder in 1985, followed by by Venzo de Sabbata and C. Sivaram in 1993. Also this physics discovery was thus made much too late. In 1995, Corrado Massa took up the idea. Independently, Ludwik Kostro in 1999, Christoph Schiller just before 2000 and Gary Gibbons in the years before 2002 arrived at the same concept. Gary Gibbons was inspired by a book by Oliver Lodge; he explains that the maximum force value follows from general relativity; he does not make a statement about the converse, nor do the other authors. The statement of maximum force as a *fundamental principle* seems original to Christoph Schiller.

The temporal order of the first papers on maximum force seems to be H.-J. Treder, *The Planckions as Largest Elementary Particles and as Smallest Test Bodies*, Foundations of Physics **15**, nr 2, 1985, followed by V. de Sabbata & C. Sivaram, *On limiting field strengths in gravitation*, Foundations of Physics Letters **6**, pp. 561–570, 1993, then by C. Massa, *Does the gravitational constant increase?*, Astrophysics and Space Science **232**, pp. 143–148, 1995, and by L. Kostro & B. Lange, *Is c^4/G the greatest possible force in nature?*, Physics Essays **12**, pp. 182–189, 1999. The next references are the paper by G. W. Gibbons, *The maximum tension principle in general relativity*, Foundations of Physics **32**, pp. 1891–1901, 2002, preprint at arxiv.org/abs/hep-th/0210109 – though he developed the ideas before that date – and the older versions of the present text, i.e., Christoph Schiller, *Motion Mountain – The Adventure of Physics*, a free pdf available at www.motionmountain.net. Then came C. Schiller, *Maximum force and minimum distance: physics in limit statements*, preprint at arxiv.org/abs/physics/0309118, and C. Schiller, *General relativity and cosmology derived from principle of maximum power or force*, International Journal of Theoretical Physics **44**, pp. 1629–1647, 2005, preprint at arxiv.org/abs/physics/0607090. See also R. Beig, G. W. Gibbons & R. M. Schoen, *Gravitating opposites attract*, Classical and Quantum Gravity **26**, p. 225013, 2009. preprint at arxiv.org/abs/0907.1193.

A detailed discussion of maximum force and power is given in the volume on general relativity, Ref. 2. Cited on pages 33, 43, 296, and 458.

Vol. II, page 107

20 Maximal luminosity is often mentioned in connection with gravitational wave detection; nevertheless, the general power maximum has never been mentioned before. See for example L. Ju, D. G. Blair & C. Zhao, *Detection of gravitational waves*, Reports on Progress in Physics **63**, pp. 1317–1427, 2000. See also C. W. Misner, K. S. Thorne & J. A. Wheeler, *Gravitation*, Freeman, 1973, page 980. Cited on page 33.

21 See for example Wolfgang Rindler, *Relativity – Special, General and Cosmological*, Oxford University Press, 2001, p. 70 ss, or Ray d'Inverno, *Introducing Einstein's Relativity*, Clarendon Press, 1992, p. 36 ss. Cited on page 34.

22 T. Jacobson, *Thermodynamics of spacetime: the Einstein equation of state*, Physical Review Letters **75**, pp. 1260–1263, 1995, preprint at arxiv.org/abs/gr-qc/9504004; this deep article remains fascinating to this day. Even the author was scared to draw all the possible conclusions. The general concepts are explained, almost without formulae, in L. Smolin, *On the nature of quantum fluctuations and their relation to gravitation and the principle of inertia*, Classical and Quantum Gravity **3**, pp. 347–359, 1986. Cited on pages 34 and 295.

23 This relation was pointed out by Achim Kempf. The story is told in A. D. Sakharov, General Relativity and Gravitation **32**, pp. 365–367, 2000, a reprint of his paper Doklady Akademii Nauk SSSR **177**, pp. 70–71, 1967. Cited on pages 35 and 44.

24 Indeterminacy relations in general relativity are discussed in C. A. Mead, *Possible connection between gravitation and fundamental length*, Physical Review B **135**, pp. 849–862, 1964. The generalized indeterminacy relation is implicit on page 852, but the issue is explained rather unclearly. Probably the author considered the result too simple to be mentioned explicitly. (That paper took 5 years to get published; comments on the story, written 37 years later, are found at C. A. Mead, *Walking the Planck length through history*, Physics Today **54**, p. 15 and p. 81, 2001, with a reply by Frank Wilczek.) See also P. K. Townsend, *Small-scale structure of space-time as the origin of the gravitational constant*, Physical Review D **15**, pp. 2795–2801, 1977, or the paper by M.-T. Jaekel & S. Renaud, *Gravitational quantum limit for length measurement*, Physics Letters A **185**, pp. 143–148, 1994. Cited on pages 36, 67, 68, 69, 72, and 120.

25 M. Kramer & al., *Tests of general relativity from timing the double pulsar*, preprint at arxiv.org/abs/astro-ph/060941. Cited on page 36.

26 Minimal length and minimal time intervals are discussed, for example, by G. Amelino-Camelia, *Limits on the measurability of space-time distances in (the semiclassical approximation of) quantum gravity*, Modern Physics Letters A **9**, pp. 3415–3422, 1994, preprint at arxiv.org/abs/gr-qc/9603014, and by Y. J. Ng & H. van Dam, *Limit to space-time measurement*, Modern Physics Letters A **9**, pp. 335–340, 1994. Many other authors have explored the topic. Cited on pages 38 and 68.

27 Maximal curvature, as well as area and volume quantization, are discussed in A. Ashtekar, *Quantum geometry and gravity: recent advances*, preprint at arxiv.org/abs/gr-qc/0112038 and in A. Ashtekar, *Quantum geometry in action: big bang and black holes*, preprint at arxiv.org/abs/math-ph/0202008. Cited on pages 38, 76, and 457.

28 Maximons, elementary particles of Planck mass, are discussed by A. D. Sakharov, *Vacuum quantum fluctuations in curved space and the theory of gravitation*, Soviet Physics – Doklady **12**, pp. 1040–1041, 1968. Cited on pages 40, 79, and 123.

29 Wolfgang Rindler, *Relativity – Special, General and Cosmological*, Oxford University Press, 2001, p. 230. Cited on page 42.

30 Several incorrect counterclaims to the entropy limit were made in R. Bousso, *The holographic principle*, Review of Modern Physics **74**, pp. 825–874, 2002, preprint at arxiv.org/abs/hep-th/0203101. However, this otherwise good review has some errors in its arguments, as explained on page 147 in volume V. Bousso has changed his position in the meantime; he now accepts the entropy limit. Cited on pages 44, 48, 455, and 457.

31 Gamma-ray bursts are discussed by G. Preparata, R. Ruffini & S. -S. Xue, *The dyadosphere of black holes and gamma-ray bursts*, Astronomy and Astrophysics **338**, pp. L87–L90, 1998, and C. L. Bianco, R. Ruffini & S. -S. Xue, *The elementary spike produced by a pure e^+e^- pair-electromagnetic pulse from a black hole: the PEM pulse*, Astronomy and Astrophysics **368**, pp. 377–390, 2001. Cited on page 45.

32 See for example the review in C.W.J. Beenakker & al., *Quantum transport in semiconductor nanostructures*, pp. 1–228, in H. Ehrenreich & D. Turnbull editors, *Solid State Physics*, volume 44, Academic Press, 1991. Cited on page 45.

33 A discussion of a different electrical indeterminacy relation, between current and charge, can be found in Y-Q. Li & B. Chen, *Quantum theory for mesoscopic electronic circuits and its applications*, preprint at arxiv.org/abs/cond-mat/9907171. Cited on page 45.

34 Hans C. Ohanian & Remo Ruffini, *Gravitation and Spacetime*, W.W. Norton & Co., 1994. Cited on pages 46 and 462.

35 The entropy limit for black holes is discussed by J. D. Bekenstein, *Entropy bounds and black hole remnants*, Physical Review D **49**, pp. 1912–1921, 1994. See also J. D. Bekenstein, *Universal upper bound on the entropy-to-energy ratio for bounded systems*, Physical Review D **23**, pp. 287–298, 1981. Cited on pages 48 and 133.

36 The statement is also called the *Kovtun-Son-Starinets conjecture*. It was published as P. Kovtun, D. T. Son & A. O. Starinets, *A viscosity bound conjecture*, preprint at arxiv.org/abs/hep-th/0405231. See also P. Kovtun, D. T. Son & A. O. Starinets, *Viscosity in strongly interacting quantum field theories from black hole physics*, Physical Review Letters **44**, p. 111601, 2005. For an experimental verification, see U. Hohm, *On the ratio of the shear viscosity to the density of entropy of the rare gases and H2, N2, CH4, and CF4*, Chemical Physics **444**, pp. 39–42, 2014. Cited on page 49.

37 Brian Greene, *The Elegant Universe – Superstrings, Hidden Dimensions, and the Quest for the Ultimate Theory*, Jonathan Cape 1999. Cited on page 53.

38 S. Weinberg, *The cosmological constant problem*, Reviews of Modern Physics **61**, pp. 1–23, 1989. Cited on page 58.

39 Steven Weinberg, *The Quantum Theory of Fields*, Cambridge University Press, volumes I, 1995, and II, 1996. Cited on page 58.

40 See the excellent presentation on the cosmological constant in general relativity by E. Bianchi & C. Rovelli, *Why all these prejudices against a constant?*, preprint at arxiv.org/abs/1002.3966 Cited on page 58.

41 The difficulties are summarised by B. S. DeWitt, *Quantum field theory in curved spacetime*, Physics Reports **19**, pp. 295–357, 1975. Cited on page 58.

42 C. W. Misner, K. S. Thorne & J. A. Wheeler, *Gravitation*, Freeman, 1973. Cited on pages 59, 60, and 68.

43 J. A. Wheeler, in *Relativity, Groups and Topology*, edited by C. DeWitt & B. S. DeWitt, Gordon and Breach, 1994. See also J. A. Wheeler, *Physics at the Planck length*, International Journal of Modern Physics A **8**, pp. 4013–4018, 1993. However, his claim that spin 1/2 *requires* topology change is *refuted* by the strand model of the vacuum. Cited on page 59.

44 J. L. Friedman & R. D. Sorkin, *Spin 1/2 from gravity*, Physical Review Letters **44**, pp. 1100–1103, 1980. Cited on page 59.

45 A. P. Balachandran, G. Bimonte, G. Marmo & A. Simoni, *Topology change and quantum physics*, Nuclear Physics B **446**, pp. 299–314, 1995, preprint at arxiv.org/abs/hep-th/9503046. Cited on page 59.

46 J. Ehlers, *Introduction – Survey of Problems*, pp. 1–10, in J. Ehlers, editor, *Sistemi gravitazionali isolati in relatività generale*, Rendiconti della scuola internazionale di fisica "Enrico Fermi", LXVII° corso, Società Italiana di Fisica/North Holland, 1979. Cited on page 59.

47 See C. Schiller, *Le vide diffère-t-il de la matière?* in E. Gunzig & S. Diner editors, *Le Vide – Univers du tout et du rien – Des physiciens et des philosophes s'interrogent*, Les Éditions de l'Université de Bruxelles, 1998. An older, English-language version is available as C. Schiller, *Does matter differ from vacuum?* preprint at arxiv.org/abs/gr-qc/9610066. Cited on pages 59, 120, 122, 123, 124, 135, and 136.

48 See for example Richard P. Feynman, Robert B. Leighton & Matthew Sands, *The Feynman Lectures on Physics*, Addison Wesley, 1977. Cited on page 60.

49 Steven Weinberg, *Gravitation and Cosmology*, Wiley, 1972. Cited on pages 60, 66, and 68.

50 The argument is given e.g. in E. P. Wigner, *Relativistic invariance and quantum phenomena*, Reviews of Modern Physics **29**, pp. 255–258, 1957. Cited on page 65.

51 The starting point for the following arguments is taken from M. Schön, *Operative time definition and principal indeterminacy*, preprint at arxiv.org/abs/gr-qc/9304024, and from T. Padmanabhan, *Limitations on the operational definition of space-time events and quantum gravity*, Classical and Quantum Gravity **4**, pp. L107–L113, 1987; see also Padmanabhan's earlier papers referenced there. Cited on page 65.

52 W. Heisenberg, *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*, Zeitschrift für Physik **43**, pp. 172–198, 1927. Cited on page 65.

53 E. H. Kennard, *Zur Quantenmechanik einfacher Bewegungstypen*, Zeitschrift für Physik **44**, pp. 326–352, 1927. Cited on page 65.

54 M. G. Raymer, *Uncertainty principle for joint measurement of noncommuting variables*, American Journal of Physics **62**, pp. 986–993, 1994. Cited on page 65.

55 H. Salecker, & E. P. Wigner, *Quantum limitations of the measurement of space-time distances*, Physical Review **109**, pp. 571–577, 1958. Cited on pages 66, 94, and 115.

56 E. J. Zimmerman, *The macroscopic nature of space-time*, American Journal of Physics **30**, pp. 97–105, 1962. Cited on pages 66, 94, and 115.

57 J. D. Bekenstein, *Black holes and entropy*, Physical Review D 7, pp. 2333–2346, 1973. Cited on pages 66, 133, and 292.

58 S. W. Hawking, *Particle creation by black holes*, Communications in Mathematical Physics **43**, pp. 199–220, 1975; see also S. W. Hawking, *Black hole thermodynamics*, Physical Review D **13**, pp. 191–197, 1976. Cited on pages 66, 133, and 292.

59 P. Gibbs, *The small scale structure of space-time: a bibliographical review*, preprint at arxiv.org/abs/hep-th/9506171. Cited on pages 66 and 85.

60 The impossibility of determining temporal ordering in quantum theory is discussed by J. Oppenheimer, B. Reznik & W. G. Unruh, *Temporal ordering in quantum mechanics*, Journal of Physics A **35**, pp. 7641–7652, 2001, preprint at arxiv.org/abs/quant-ph/0003130. Cited on page 67.

61 M. -T. Jaekel & S. Renaud, *Gravitational quantum limit for length measurement*, Physics Letters A **185**, pp. 143–148, 1994. Cited on page 68.

62 D. V. Ahluwalia, *Quantum measurement, gravitation and locality*, Physics Letters B **339**, pp. 301–303, 1994, preprint at arxiv.org/abs/gr-qc/9308007. Cited on page 68.

63 L. Garay, *Quantum gravity and minimum length*, International Journal of Modern Physics A **10**, pp. 145–165, 1995, preprint at arxiv.org/abs/gr-qc/9403008. This paper also includes an extensive bibliography. See also R. J. Adler & D. I. Santiago, *On gravity and the uncertainty principle*, Modern Physics Letters A **14**, pp. 1371–1381, 1999, preprint at arxiv.org/abs/gr-qc/9904026. Cited on page 68.

64 C. Rovelli & L. Smolin, *Discreteness of area and volume in quantum gravity*, Nuclear Physics B **442**, pp. 593–619, 1995. R. Loll, *The volume operator in discretized quantum gravity*, preprint at arxiv.org/abs/gr-qc/9506014. See also C. Rovelli, *Notes for a brief history of quantum gravity*, preprint at arxiv.org/abs/gr-qc/0006061. Cited on page 69.

65 D. Amati, M. Ciafaloni & G. Veneziano, *Superstring collisions at Planckian energies*, Physics Letters B **197**, pp. 81–88, 1987. D. J. Gross & P. F. Mende, *The high energy behavior of string scattering amplitudes*, Physics Letters B **197**, pp. 129–134, 1987. K. Konishi, G. Paffuti & P. Provero, *Minimum physical length and the generalized uncertainty principle*, Physics Letters B **234**, pp. 276–284, 1990. P. Aspinwall, *Minimum distances in non-trivial string target spaces*, Nuclear Physics B **431**, pp. 78–96, 1994, preprint at arxiv.org/abs/hep-th/9404060. Cited on page 69.

66 M. Maggiore, *A generalised uncertainty principle in quantum mechanics*, Physics Letters B **304**, pp. 65–69, 1993. Cited on page 69.

67 A simple approach is S. Doplicher, K. Fredenhagen & J. E. Roberts, *Space-time quantization induced by classical gravity*, Physics Letters B **331**, pp. 39–44, 1994. Cited on pages 69 and 84.

68 A. Kempf, *Uncertainty relation in quantum mechanics with quantum group symmetry*, Journal of Mathematical Physics 35, pp. 4483–4496, 1994. A. Kempf, *Quantum groups* *and quantum field theory with nonzero minimal uncertainties in positions and momenta*, Czechoslovak Journal of Physics **44**, pp. 1041–1048, 1994. Cited on page 69.

69 E. J. Hellund & K. Tanaka, *Quantized space-time*, Physical Review **94**, pp. 192–195, 1954. Cited on page 70.

70 This intriguing extract from a letter by Einstein was made widely known by John J. Stachel, in his paper *The other Einstein: Einstein contra field theory*, that is best found in his book *Einstein from 'B' to 'Z'*, Birkhäuser, 2002. The German original of the letter is found in Robert Schulmann, A. J. Knox, Michel Janssen & József Illy, *The Collected Papers of Albert Einstein, Volume 8A – The Berlin Years: Correspondence, 1914–1917*, letter 299, Princeton University Press, 1998. Barbara Wolff helped in clarifying several details in the German original. The letter is now available online, at einsteinpapers.press.princeton.edu/vol8a-doc/463. Cited on page 70.

71 A. Peres & N. Rosen, *Quantum limitations on the measurement of gravitational fields*, Physical Review **118**, pp. 335–336, 1960. Cited on page 72.

72 It is the first definition in Euclid's *Elements*, *c.* 300 bce. For an English translation see T. Heath, *The Thirteen Books of the Elements*, Dover, 1969. Cited on page 73.

73 A beautiful description of the Banach–Tarski paradox is the one by Ian Stewart, *Paradox of the spheres*, New Scientist, 14 January 1995, pp. 28–31. Cited on page 73.

74 H. S. Snyder, *Quantized space-time*, Physical Review **71**, pp. 38–41, 1947. H. S. Snyder, *The electromagnetic field in quantized space-time*, Physical Review **72**, pp. 68–74, 1947. A. Schild, *Discrete space-time and integral Lorentz transformations*, Physical Review **73**, pp. 414–415, 1948. E. L. Hill, *Relativistic theory of discrete momentum space and discrete space-time*, Physical Review **100**, pp. 1780–1783, 1950. H. T. Flint, *The quantization of space-time*, Physical Review **74**, pp. 209–210, 1948. A. Das, *Cellular space-time and quantum field theory*, Il Nuovo Cimento **18**, pp. 482–504, 1960. Cited on page 75.

75 D. Finkelstein, *'Superconducting' causal nets*, International Journal of Theoretical Physics **27**, pp. 473–519, 1985. Cited on page 75.

76 N. H. Christ, R. Friedberg & T. D. Lee, *Random lattice field theory: general formulation*, Nuclear Physics B **202**, pp. 89–125, 1982. G. 't Hooft, *Quantum field theory for elementary particles – is quantum field theory a theory?*, Physics Reports **104**, pp. 129–142, 1984. Cited on page 75.

77 For a discussion, see R. Sorabji, *Time, Creation and the Continuum: Theories in Antiquity and the Early Middle Ages*, Duckworth, 1983. Cited on page 75.

78 See, for example, L. Bombelli, J. Lee, D. Meyer & R. D. Sorkin, *Space-time as a causal set*, Physical Review Letters **59**, pp. 521–524, 1987. G. Brightwell & R. Gregory, *Structure of random space-time*, Physical Review Letters **66**, pp. 260–263, 1991. Cited on page 75.

79 The false belief that particles like quarks or electrons are composite is slow to die out. See for example: S. Fredriksson, *Preon prophecies by the standard model*, preprint at arxiv.org/abs/hep-ph/0309213. Preon models gained popularity in the 1970s and 1980s, in particular through the papers by J. C. Pati & A. Salam, *Lepton number as the fourth "color"*, Physical Review D **10**, pp. 275–289, 1974, H. Harari, *A schematic model of quarks and leptons*, Physics Letters B **86**, pp. 83–86, 1979, M. A. Shupe, *A composite model of leptons and quarks*, Physics Letters B **86**, pp. 87–92, 1979, and H. Fritzsch & G. Mandelbaum, *Weak interactions as manifestations of the substructure of leptons and quarks*, Physics Letters B **102**, pp. 319–322, 1981. Cited on page 77.

80 N. F. Ramsey & A. Weis, *Suche nach permanenten elektrischen Dipolmomenten: ein Test der Zeitumkehrinvarianz*, Physikalische Blätter **52**, pp. 859–863, 1996. See also W. Bernreuther & M. Suzuki, *The electric dipole moment of the electron*, Reviews of Modern Physics **63**, pp. 313–340, 1991, and the musings in Hans Dehmelt, *Is the electron a composite particle?*, Hyperfine Interactions **81**, pp. 1–3, 1993. Cited on page 78.

81 K. Akama, T. Hattori & K. Katsuura, *Naturalness bounds on dipole moments from new physics*, preprint at arxiv.org/abs/hep-ph/0111238. Cited on page 78.

82 The paper by J. Baron & al., *Order of magnitude smaller limit on the electric dipole moment of the electron*, preprint at arxiv.org/abs/1310.7534 gives an upper experimental limit to the dipole moment of the electron of $8.7 \cdot 10^{-31}\, e$ m. Cited on page 78.

83 C. Wolf, *Upper limit for the mass of an elementary particle due to discrete time quantum mechanics*, Il Nuovo Cimento B **109**, pp. 213–218, 1994. Cited on page 80.

84 W. G. Unruh, *Notes on black hole evaporation*, Physical Review D **14**, pp. 870–875, 1976. W. G. Unruh & R. M. Wald, *What happens when an accelerating observer detects a Rindler particle*, Physical Review D **29**, pp. 1047–1056, 1984. Cited on page 82.

85 The first example was J. Magueijo & L. Smolin, *Lorentz invariance with an invariant energy scale*, Physical Review Letters **88**, p. 190403, 2002, preprint at arxiv.org/abs/hep-th/0112090. They propose a modification of the mass energy relation of the kind

$$E = \frac{c^2 \gamma m}{1 + \frac{c^2 \gamma m}{E_{\rm Pl}}} \quad \text{and} \quad p = \frac{\gamma m v}{1 + \frac{c^2 \gamma m}{E_{\rm Pl}}} . \qquad (215)$$

Another, similar approach of recent years, with a different proposal, is called 'doubly special relativity'. A recent summary is G. Amelino-Camelia, *Doubly-special relativity: first results and key open problems*, International Journal of Modern Physics **11**, pp. 1643–1669, 2002, preprint at arxiv.org/abs/gr-qc/0210063. The paper shows how conceptual problems hinder the advance of the field. Another such discussion R. Aloisio, A. Galante, A. F. Grillo, E. Luzio & F. Méndez, *Approaching space-time through velocity in doubly special relativity*, preprint at arxiv.org/abs/gr-qc/0410020. The lesson from these attempts is simple: special relativity *cannot* be modified to include a limit energy without also including general relativity and quantum theory. Cited on pages 84 and 282.

86 W. Jauch, *Heisenberg's uncertainty relation and thermal vibrations in crystals*, American Journal of Physics **61**, pp. 929–932, 1993. Cited on page 84.

87 H. D. Zeh, *On the interpretation of measurement in quantum theory*, Foundations of Physics **1**, pp. 69–76, 1970. Cited on page 85.

88 See Y. J. Ng, W. A. Christiansen & H. van Dam, *Probing Planck-scale physics with extragalactic sources?*, Astrophysical Journal **591**, pp. L87–L90, 2003, preprint at arxiv.org/abs/astro-ph/0302372; D. H. Coule, *Planck scale still safe from stellar images*, Classical and Quantum Gravity **20**, pp. 3107–3112, 2003, preprint at arxiv.org/abs/astro-ph/0302333. Negative experimental results (and not always correct calculations) are found in R. Lieu & L. Hillman, *The phase coherence of light from extragalactic sources – direct evidence against first order Planck scale fluctuations in time and space*, Astrophysical Journal **585**, pp. L77–L80, 2003, and R. Ragazzoni, M. Turatto & W. Gaessler, *The lack of observational evidence for the quantum structure of spacetime at Planck scales*, Astrophysical Journal **587**, pp. L1–L4, 2003. Cited on page 88.

89 B. E. Schaefer, *Severe limits on variations of the speed of light with frequency*, Physical Review Letters **82**, pp. 4964–4966, 21 June 1999. Cited on page 88.

90 A.A. Abdo & al., (Fermi GBM/LAT collaborations) *Testing Einstein's special relativity with Fermi's short hard gamma-ray burst GRB090510*, preprint at arxiv.org/0908.1832. Cited on page 88.

91 G. Amelino-Camelia, J. Ellis, N.E. Mavromatos, D.V. Nanopoulos & S. Sakar, *Potential sensitivity of gamma-ray-burster observations to wave dispersion in vacuo*, Nature **393**, pp. 763–765, 1998, preprint at arxiv.org/abs/astro-ph/9712103. Cited on page 88.

92 G. Amelino-Camelia, *Phenomenological description of space-time foam*, preprint at arxiv.org/abs/gr-qc/0104005. The paper includes a clearly written overview of present experimental approaches to detecting quantum gravity effects. See also his update G. Amelino-Camelia, *Quantum-gravity phenomenology: status and prospects*, preprint at arxiv.org/abs/gr-qc/0204051. Cited on pages 88 and 89.

93 G. Amelino-Camelia, *An interferometric gravitational wave detector as a quantum gravity apparatus*, Nature **398**, pp. 216–218, 1999, preprint at arxiv.org/abs/gr-qc/9808029. Cited on page 88.

94 F. Karolyhazy, *Gravitation and quantum mechanics of macroscopic objects*, Il Nuovo Cimento **A42**, pp. 390–402, 1966. Y. J. Ng & H. van Dam, *Limit to space-time measurement*, Modern Physics Letters A **9**, pp. 335–340, 1994. Y. J. Ng & H. van Dam, Modern Physics Letters A *Remarks on gravitational sources*, **10**, pp. 2801–2808, 1995. The discussion is neatly summarised in Y. J. Ng & H. van Dam, *Comment on 'Uncertainty in measurements of distance'*, preprint at arxiv.org/abs/gr-qc/0209021. See also Y. J. Ng, *Spacetime foam*, preprint at arxiv.org/abs/gr-qc/0201022. Cited on pages 88 and 94.

95 L. J. Garay, *Spacetime foam as a quantum thermal bath*, Physics Review Letters **80**, pp. 2508–2511, 1998, preprint at arxiv.org/abs/gr-qc/9801024. Cited on page 89.

96 G. Amelino-Camelia & T. Piran, *Planck-scale deformation of Lorentz symmetry as a solution to the UHECR and the TeV-γ paradoxes*, preprint at arxiv.org/astro-ph/0008107, 2000. Cited on page 89.

97 R. P. Woodard, *How far are we from the quantum theory of gravity?*, preprint at arxiv.org/abs/0907.4238. For a different point of view, see L. Smolin, *Generic predictions of quantum theories of gravity*, preprint at arxiv.org/abs/hep-th/0605052. Cited on pages 89 and 306.

98 A similar point of view, often called monism, was proposed by Baruch Spinoza, *Ethics Demonstrated in Geometrical Order*, 1677, originally in Latin; an affordable French edition is Baruch Spinoza, *L'Ethique*, Folio-Gallimard, 1954. For a discussion of his ideas, especially his monism, see Don Garret editor, *The Cambridge Companion to Spinoza*, Cambridge University Press, 1996, or any general text on the history of philosophy. Cited on page 89.

99 See the lucid discussion by G. F. R. Ellis & T. Rothman, *Lost horizons*, American Journal of Physics **61**, pp. 883–893, 1993. Cited on pages 94, 98, and 99.

100 See, for example, the Hollywood film *Contact* by Robert Zemeckis, based on the book by Carl Sagan, *Contact*, Simon & Schuster, 1985. Cited on page 100.

101 See, for example, the international bestseller by Stephen Hawking, *A Brief History of Time – From the Big Bang to Black Holes*, 1988. Cited on page 102.

102 L. Rosenfeld, *Quantentheorie und Gravitation*, in H.-J. Treder, editor, *Entstehung, Entwicklung und Perspektiven der Einsteinschen Gravitationstheorie*, Springer Verlag, 1966. Cited on page 105.

103 Holography in high-energy physics is connected with the work of 't Hooft and Susskind. See for example G. 't Hooft, *Dimensional reduction in quantum gravity*, pp. 284–296, in A. Ali, J. Ellis & S. Randjbar-Daemi, *Salaamfeest*, 1993, or the much-cited paper by L. Susskind, *The world as a hologram*, Journal of Mathematical Physics **36**, pp. 6377–6396, 1995, preprint at arxiv.org/abs/hep-th/9409089. A good modern overview is Ref. 30. Cited on pages 106 and 114.

104 D. Bohm & B. J. Hiley, *On the intuitive understanding of nonlocality as implied by quantum theory*, Foundations of Physics 5, pp. 93–109, 1975. Cited on page 107.

105 S. Lloyd, *Computational capacity of the universe*, Physical Review Letters **88**, p. 237901, 2002. Cited on page 109.

106 Gottfried Wilhelm Leibniz, *La Monadologie*, 1714. Written in French, it is available freely at www.uqac.uquebec.ca/zone30/Classiques_des_sciences_sociales and in various other languages on other websites. Cited on page 109.

107 See, for example, H. Wussing & P. S. Alexandrov editors, *Die Hilbertschen Probleme*, Akademische Verlagsgesellschaft Geest & Portig, 1983, or Ben H. Yandell, *The Honours Class: Hilbert's Problems and their Solvers*, A.K. Peters, 2002. Cited on page 110.

108 A large part of the study of dualities in string and M theory can be seen as investigations into the detailed consequences of extremal identity. For a review of dualities, see P. C. Argyres, *Dualities in supersymmetric field theories*, Nuclear Physics Proceedings Supplement **61**, pp. 149–157, 1998, preprint at arxiv.org/abs/hep-th/9705076. A classical version of duality is discussed by M. C. B. Abdalla, A. L. Gadelka & I. V. Vancea, *Duality between coordinates and the Dirac field*, preprint at arxiv.org/abs/hep-th/0002217. Cited on page 114.

109 See L. Susskind & J. Uglum, *Black holes, interactions, and strings*, preprint at arxiv.org/abs/hep-th/9410074, or L. Susskind, *String theory and the principle of black hole complementarity*, Physical Review Letters **71**, pp. 2367–2368, 1993, and M. Karliner, I. Klebanov & L. Susskind, *Size and shape of strings*, International Journal of Modern Physics A **3**, pp. 1981–1996, 1988, as well as L. Susskind, *Structure of hadrons implied by duality*, Physical Review D **1**, pp. 1182–1186, 1970. Cited on pages 119 and 134.

110 M. Planck, *Über irreversible Strahlungsvorgänge*, Sitzungsberichte der königlich-preußischen Akademie der Wissenschaften zu Berlin pp. 440–480, 1899. Today it is commonplace to use Dirac's $\hbar = h/2\pi$ instead of Planck's h, which Planck originally called b. Cited on page 120.

111 P. Facchi & S. Pascazio, *Quantum Zeno and inverse quantum Zeno effects*, pp. 147–217, in E. Wolf editor, *Progress in Optics*, **42**, 2001. Cited on page 123.

112 Aristotle, *Of Generation and Corruption*, book I, part 2. See Jean-Paul Dumont, *Les écoles présocratiques*, Folio Essais, Gallimard, p. 427, 1991. Cited on page 123.

113 See for example the speculative model of vacuum as composed of Planck-size spheres proposed by F. Winterberg, Zeitschrift für Naturforschung **52a**, p. 183, 1997. Cited on page 124.

114 The Greek salt-and-water argument and the fish argument are given by Lucrece, in full Titus Lucretius Carus, *De natura rerum*, *c.* 60 bce. Cited on pages 125 and 140.

115 J. H. Schwarz, *The second superstring revolution*, Colloquium-level lecture presented at the Sakharov Conference in Moscow, May 1996, preprint at arxiv.org/abs/hep-th/9607067. Cited on pages 126 and 128.

116 Simplicius, *Commentary on the Physics of Aristotle*, 140, 34. This text is cited in Jean-Paul Dumont, *Les écoles présocratiques*, Folio Essais, Gallimard, p. 379, 1991. Cited on page 127.

117 D. Olive & C. Montonen, *Magnetic monopoles as gauge particles*, Physics Letters **72B**, pp. 117–120, 1977. Cited on page 128.

118 A famous fragment from Diogenes Laërtius (IX 72) quotes Democritus as follows: 'By convention hot, by convention cold, but in reality, atoms and void; and also in reality we know nothing, since truth is at the bottom.' Cited on page 129.

119 This famous statement is found at the beginning of chapter XI, 'The Physical Universe', in Arthur Eddington, *The Philosophy of Physical Science*, Cambridge, 1939. Cited on page 130.

120 Plato, *Parmenides*, *c.* 370 bce. It has been translated into most languages. Reading it aloud, like a song, is a beautiful experience. A pale reflection of these ideas is Bohm's concept of 'unbroken wholeness'. Cited on page 130.

121 P. Gibbs, *Event-symmetric physics*, preprint at arxiv.org/abs/hep-th/9505089; see also his website www.weburbia.com/pg/contents.htm. Cited on page 130.

122 J. B. Hartle, & S. W. Hawking, *Path integral derivation of black hole radiance*, Physical Review D **13**, pp. 2188–2203, 1976. See also A. Strominger & C. Vafa, *Microscopic origin of Bekenstein–Hawking entropy*, Physics Letters B **379**, pp. 99–104, 1996, preprint at arxiv.org/abs/hep-th/9601029. For another derivation of black hole entropy, see G. T. Horowitz & J. Polchinski, *A correspondence principle for black holes and strings*, Physical Review D 55, pp. 6189–6197, 1997, preprint at arxiv.org/abs/hep-th/9612146. Cited on pages 133 and 142.

123 J. Maddox, *When entropy does not seem extensive*, Nature **365**, p. 103, 1993. The issue is now explored in all textbooks discussing black holes. John Maddox (b. 1925 Penllergaer, d. 1999 Abergavenny) was famous for being one of the few people who was knowledgeable in most natural sciences. Cited on page 133.

124 L. Bombelli, R. K. Koul, J. Lee & R. D. Sorkin, *Quantum source of entropy of black holes*, Physical Review D **34**, pp. 373–383, 1986. Cited on page 133.

125 The analogy between polymers and black holes is due to G. Weber, *Thermodynamics at boundaries*, Nature **365**, p. 792, 1993. Cited on page 133.

126 See the classic text by Pierre-Gilles de Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, 1979. Cited on page 134.

127 See for example S. Majid, *Introduction to braided geometry and q-Minkowski space*, preprint at arxiv.org/abs/hep-th/9410241, or S. Majid, *Duality principle and braided geometry*, preprint at arxiv.org/abs/hep-th/9409057. Cited on pages 135 and 136.

128 The relation between spin and statistics has been studied recently by M. V. Berry & J. M. Robbins, *Quantum indistinguishability: spin–statistics without relativity or field theory?*, in R. C. Hilborn & G. M. Tino editors, *Spin–Statistics Connection and Commutation Relations*, American Institute of Physics, 2000. Cited on page 137.

129 A. Gregori, *Entropy, string theory, and our world*, preprint at arxiv.org/abs/hep-th/0207195. Cited on pages 138 and 139.

130 String cosmology is a pastime for many. Examples include N. E. Mavromatos, *String cosmology*, preprint at arxiv.org/abs/hep-th/0111275, and N. G. Sanchez, *New developments in string gravity and string cosmology – a summary report*, preprint at arxiv.org/abs/hep-th/0209016. Cited on page 139.

131 On the present record, see en.wkipedia.org/wiki/Ultra-high-energy_cosmic_ray and fr.wkipedia.org/wiki/Zetta-particule. Cited on page 140.

132 P. F. Mende, *String theory at short distance and the principle of equivalence*, preprint at arxiv.org/abs/hep-th/9210001. Cited on page 140.

133 An example is given by A. A. Slavnov, *Fermi–Bose duality via extra dimension*, preprint at arxiv.org/abs/hep-th/9512101. See also the standard work by Michael Stone editor, *Bosonization*, World Scientific, 1994. Cited on page 140.

134 A weave model of space-time appears in certain approaches to quantum gravity, such as Ref. 27. On a slightly different topic, see also S. A. Major, *A spin network primer*, preprint at arxiv.org/abs/gr-qc/9905020. Cited on page 140.

135 L. Smolin & Y. Wan, *Propagation and interaction of chiral states in quantum gravity*, preprint at arxiv.org/abs/0710.1548, and references therein. Cited on page 140.

136 A good introduction into his work is the paper D. Kreimer, *New mathematical structures in renormalisable quantum field theories*, Annals of Physics **303**, pp. 179–202, 2003, erratum ibid. 305, p. 79, 2003, preprint at arxiv.org/abs/hep-th/0211136. Cited on page 141.

137 Introductions to holography include E. Alvarez, J. Conde & L. Hernandez, *Rudiments of holography*, preprint at arxiv.org/abs/hep-th/0205075, and Ref. 30. The importance of holography in theoretical high-energy physics was underlined by the discovery of J. Maldacena, *The large N limit of superconformal field theories and supergravity*, preprint at arxiv.org/abs/hep-th/9711200. Cited on page 141.

138 X.-G. Wen, *From new states of matter to a unification of light and electrons*, preprint at arxiv.org/abs/0508020. Cited on page 141.

139 J. S. Avrin, *A visualizable representation of the elementary particles*, Journal of Knot Theory and Its Ramifications **14**, pp. 131–176, 2005. Cited on pages 141 and 347.

140 The well-known ribbon model is presented in S. Bilson-Thompson, *A topological model of composite preons*, preprint at arxiv.org/hep-ph/0503213; S. Bilson-Thompson, F. Markopoulou & L. Smolin, *Quantum gravity and the standard model*, preprint at arxiv.org/hep-th/0603022; S. Bilson-Thompson, J. Hackett, L. Kauffman & L. Smolin, *Particle identifications from symmetries of braided ribbon network invariants*, preprint at arxiv.org/abs/0804.0037; S. Bilson-Thompson, J. Hackett & L. Kauffman, *Particle topology, braids, and braided belts*, preprint at arxiv.org/abs/0903.1376. Cited on pages 141, 168, and 347.

141 R. J. Finkelstein, *A field theory of knotted solitons*, preprint at arxiv.org/abs/hep-th/0701124. See also R. J. Finkelstein, *Trefoil solitons, elementary fermions, and* $SU_q(2)$, preprint at arxiv.org/abs/hep-th/0602098, R. J. Finkelstein & A. C. Cadavid, *Masses and interactions of q-fermionic knots*, preprint at arxiv.org/abs/hep-th/0507022, and R. J. Finkelstein, *A knot model suggested by the standard electroweak theory*, preprint at arxiv.org/abs/hep-th/0408218. Cited on pages 141 and 347.

142 Louis H. Kauffman, *Knots and Physics*, World Scientific, 1991. A wonderful book. Cited on pages 141 and 277.

143 S. K. Ng, *On a knot model of the* π^+ *meson*, preprint at arxiv.org/abs/hep-th/0210024, and S. K. Ng, *On a classification of mesons*, preprint at arxiv.org/abs/hep-ph/0212334. Cited on pages 141 and 347.

144 For a good introduction to superstrings, see the lectures by B. Zwiebach, *String theory for pedestrians*, agenda.cern.ch/fullAgenda.php?ida=a063319. For an old introduction to superstrings, see the famous text by M. B. Green, J. H. Schwarz & E. Witten, *Superstring Theory*, Cambridge University Press, volumes 1 and 2, 1987. Like all the other books on superstrings, they contain no statement that is applicable to or agrees with the strand model. Cited on pages 141 and 348.

145 See A. Sen, *An introduction to duality symmetries in string theory*, in *Les Houches Summer School: Unity from Duality: Gravity, Gauge Theory and Strings (Les Houches, France, 2001)*, Springer Verlag, **76**, pp. 241–322, 2002. Cited on page 141.

146 Brian Greene regularly uses the name *string conjecture*. For example, he did so in a podium discussion at TED in 2009; the video of the podium discussion can be downloaded at www.ted.org. Cited on page 142.

147 L. Susskind, *Some speculations about black hole entropy in string theory*, preprint at arxiv.org/abs/hep-th/9309145. G. T. Horowitz & J. Polchinski, *A correspondence principle for black holes and strings*, Physical Review D 55, pp. 6189–6197, 1997, preprint at arxiv.org/abs/hep-th/9612146. Cited on pages 142 and 463.

148 F. Wilczek, *Getting its from bits*, Nature **397**, pp. 303–306, 1999. Cited on page 143.

149 M. R. Douglas, *Understanding the landscape*, preprint at arxiv.org/abs/hep-th/0602266; his earlier papers also make the point. For the larger estimate, see W. Taylor & Y.-N. Wang, *The F-theory geometry with most flux vacua*, preprint at arxiv.org/abs/1511.03209. Cited on page 144.

150 The difficulties of the string conjecture are discussed in the well-known internet blog by Peter Woit, *Not even wrong*, at www.math.columbia.edu/~woit/blog. Several Nobel Prize winners for particle physics dismiss the string conjecture: Martin Veltman, Sheldon Glashow, Burton Richter, Richard Feynman and since 2009 also Steven Weinberg are among those who did so publicly. Cited on pages 144 and 168.

151 The present volume was originally started with the aim to clarify the basic principles of string theory and to simplify it as much as possible. In particular, the first six chapters and the last chapter were conceived, structured and written with that aim. They are older than the strand model. Later on, the project took an unexpected direction, as explained in Ref. 19. Cited on page 144.

152 Searches for background-free approaches are described by E. Witten, *Quantum background independence in string theory*, preprint at arxiv.org/abs/hep-th/9306122 and E. Witten, *On background-independent open string theory*, preprint at arxiv.org/abs/hep-th/9208027. Cited on page 145.

153 In fact, no other candidate model that fulfils all requirements for the unified theory is available in the literature so far. This might change in the future, though. Cited on page 151.

154 In December 2011 already I did not recall when I conceived and first explored the fundamental principle. It was between 2001 and 2006. It was a gradual process started by the exploration of the belt trick and driven by the aim to describe space in a way that produces Lorentz invariance based on extended entities (as I used to call them back then). The idea and the graphs of extended constituents for spin 1/2 is part of my volume 4 (on spin and particle exchange) since the 1990s. For many years I searched for a way to extend that model to include vacuum and the rest of physics. This was my pastime when taking the subway to work during those years. Cited on page 158.

155 S. Carlip, *The small scale structure of spacetime*, preprint at arxiv.org/abs/1009.1136. This paper deduces the existence of fluctuating lines in vacuum from a number of arguments that are completely independent of the strand model. Steven Carlip has dedicated much of his research to the exploration of this topic. One summary is S. Carlip, *Spontaneous dimensional reduction in quantum gravity*, preprint at arxiv.org/abs/1605.05694; its is also instructive to read his review S. Carlip, *Dimension and dimensional reduction in quantum gravity*, Classical and Quantum Gravity **34**, p. 193001, 2017, preprint at arxiv.org/abs/1705.05417. With the strand model in the back of one's mind, these results are even more fascinating. Cited on pages 163 and 302.

156 David Deutsch states that any good explanation must be 'hard to vary'. This must also apply to a unified model, as it claims to explain everything that is observed. See D. Deutsch, *A new way to explain explanation*, video talk at www.ted.org. Cited on pages 167 and 423.

157 L. Bombelli, J. Lee, D. Meyer & R. D. Sorkin, *Space-time as a causal set*, Physical Review Letters **59**, pp. 521–524, 1987. See also the review by J. Henson, *The causal set approach to quantum gravity*, preprint at arxiv.org/abs/gr-qc/0601121. Cited on pages 168 and 301.

158 D. Finkelstein, *Homotopy approach to quantum gravity*, International Journal of Theoretical Physics 47, pp. 534–552, 2008. Cited on pages 168 and 301.

159 L. H. Kauffman & S. J. Lomonaco, *Quantum knots*, preprint at arxiv.org/abs/quant-ph/0403228. See also S. J. Lomonaco & L. H. Kauffman, *Quantum knots and mosaics*, preprint at arxiv.org/abs/0805.0339. Cited on page 168.

160 Immanuel Kant, *Critik der reinen Vernunft*, 1781, is a famous but long book that every philospher pretends to have read. In his book, Kant introduced the 'a priori' existence of space and time. Cited on page 170.

161 The literature on circularity is rare. For two interesting exceptions, see L. H. Kauffman, *Knot logic*, downloadable from www2.math.uic.edu/~kauffman, and L. H. Kauffman, *Reflexivity and eigenform*, Constructivist Foundations **4**, pp. 121–137, 2009. Cited on page 171.

162 Information on the belt trick is scattered across many books and few papers. The best source of information on this topic are websites. For belt trick visualizations see www.evl.uic.edu/hypercomplex/html/dirac.html, www.evl.uic.edu/hypercomplex/html/handshake.html, or www.gregegan.net/APPLETS/21/21.html. For an excellent literature summary and more movies, see www.math.utah.edu/~palais/links.html. None of these sites or the cited references seem to mention that there are *many* ways to perform the belt trick; this seems to be hidden knowledge. In September 2009, Greg Egan took up my suggestion and changed his applet to show an additional version of the belt trick. Cited on pages 178 and 180.

163 There is an interesting exploration behind this analogy between a non-dissipative system – a free quantum particle moving in vacuum – and a dissipative system – a macroscopic body drawn through a viscous liquid, say honey. The first question is to discover why this analogy is possible at all. (A careful distinction between the cases with spin 0, spin 1 and spin 1/2 are necessary.) The second question is the exploration of the motion of bodies of general shape in viscous fluids at low Reynolds numbers and under constant force. For the best overview of this question, see the beautiful article by O. Gonzalez, A. B. A. Graf & J. H. Maddocks, *Dynamics of a rigid body in a Stokes fluid*, Journal of Fluid Mechanics **519**, pp. 133–160, 2004. Cited on pages 197 and 362.

Challenge 229 e

164 D. Bohm, R. Schiller & J. Tiomno, *A causal interpretation of the Pauli equation (A)*, Supplementi al Nuovo Cimento **1**, pp. 48 – 66, 1955, and D. Bohm & R. Schiller, *A causal interpretation of the Pauli equation (B)*, Supplementi al Nuovo Cimento **1**, pp. 67–91, 1955. The authors explore an unusual way to interpret the wavefunction, which is of little interest here; but doing so, they give and explore the description of Pauli spinors in terms of Euler angles. Cited on page 200.

165 Richard P. Feynman, *QED – The Strange Theory of Light and Matter*, Princeton University Press 1988. This is one of the best summaries of quantum theory ever written. Every physicist should read it. Cited on pages 201, 207, 219, 225, and 460.

166 S. Kochen & E. P. Specker, *The problem of hidden variables in quantum mechanics*, 17, pp. 59–87, 1967. This is a classic paper. Cited on page 205.

167 A. Aspect, J. Dalibard & G. Roger, *Experimental tests of Bell's inequalities using time-varying analyzers*, Physical Review Letters **49**, pp. 1804–1807, 1982, Cited on page 209.

168 L. Kauffman, *New invariants of knot theory*, American Mathematical Monthly **95**, pp. 195–242, 1987. See also the image at the start of chapter 6 of Louis H. Kauffman, *On Knots*, Princeton University Press, 1987. Cited on page 211.

169 The details on the speed of photons are explained in any textbook on quantum electrodynamics. The issue is also explained by Feynman in Ref. 165 on page 89. Cited on page 214.

170 J.-M. Lévy-Leblond, *Nonrelativitic particles and wave equations*, Communications in Mathematical Physics **6**, pp. 286–311, 1967. See also A. Galindo & C. Sánchez del Río, *Intrinsic magnetic moment as a nonrelativistic phenomenon*, American Journal of Physics **29**, pp. 582–584, 1961, and V. I. Fushchich, A. G. Nikitin & V. A. Salogub, *On the non-relativistic motion equations in the Hamiltonian form*, Reports on Mathematical Physics **13**, pp. 175–185, 1978. Cited on page 216.

171 L. Lerner, *Derivation of the Dirac equation from a relativistic representation of spin*, European Journal of Phyics **17**, pp. 172–175, 1996. Cited on page 216.

172 E. P. Battey-Pratt & T. J. Racey, *Geometric model for fundamental particles*, International Journal of Theoretical Physics **19**, pp. 437–475, 1980. Without knowing this work, C. Schiller had deduced the same results in 2008. Cited on pages 217, 218, and 470.

173 A. Abraham, *Prinzipien der Dynamik des Elektrons*, Annalen der Physik **10**, pp. 105–179, 1903, J. Frenkel, *Die Elektrodynamik des rotierenden Elektrons*, Zeitschrift für Physik **37**, pp. 243–262, 1926, L. H. Thomas, *The motion of a spinning electron*, Nature April 10, p. 514, 1926, and L. H. Thomas, *The kinematics of an electron with an axis*, Philosophical Magazine **3**, pp. 1–22, 1927. See also W. E. Baylis, *Surprising symmetries in relativistic charge dynamics*, preprint at arxiv.org/abs/physics/0410197. See also W. E. Baylis, *Quantum/classical interface: a geometric approach from the classical side*, pp. 127–154 and W. E. Baylis, *Geometry of paravector space with applications to relativistic physics*, pp. 363–387 in *Computational Noncommutative Algebra and Applications*, Proceedings of the NATO Advanced Study Institute, NATO Science Series II, vol. 136, ed. J. Byrnes, Kluwer Academic 2004. W. E. Baylis, R. Cabrera & D. Keselica, *Quantum/classical interface: fermion spin*, preprint at arxiv.org/abs/0710.3144. D. Hestenes, *Zitterbewegung Modelling*, Foundations of Physics **23**, pp. 365–386, 1993. D. Hestenes, *Zitterbewegung in quantum mechanics – a research program*, preprint at arxiv.org/abs/0802.2728. See also D. Hestenes, *Reading the electron clock*, preprint at arxiv.org/abs/0802.3227 and his webpage modelingnts.la.asu.edu/html/GAinQM.html. A. Loinger & A. Sparzani, *Dirac equation without Dirac matrices*, Il Nuovo Cimento **39**, pp. 1140–1145, 1965. D. Bohm, P. Hillion, T. Takabayasi & J.-P. Vigier, *Relativistic rotators and bilocal theory*, Progress of Theoretical Physics **23**, pp. 496–511, 1960. A. Challinor, A. Lasenby, S. Gill & C. Doran, *A relativistic, causal account of a spin measurement*, Physics Letters A **218**, pp. 128–138, 1996. E. Santamato, *The role of Dirac equation in the classical mechanics of the relativistic top*, preprint at arxiv.org/abs/0808.3237. Cited on page 219.

174 The concept of Zitterbewegung was formulated in E. Schrödinger, *Über die kräftefreie Bewegung in der relativistischen Quantenmechanik*, Berliner Berichte pp. 418–428, 1930, and *Zur Quantendynamik des Elektrons*, Berliner Berichte pp. 63–72, 1931. Numerous subsequent papers discuss these publications. Cited on page 219.

175 See for example the book by Martin Rivas, *Kinematic Theory of Spinning Particles*, Springer, 2001. Cited on page 219.

176 The basic papers in the field of stochastic quantization are W. Weizel, *Ableitung der Quantentheorie aus einem klassischen, kausal determinierten Modell*, Zeitschrift für Physik A **134**, pp. 264–285, 1953, W. Weizel, *Ableitung der Quantentheorie aus einem klassischen Modell – II*, Zeitschrift für Physik A **135**, pp. 270–273, 1954, W. Weizel, *Ableitung der quantenmechanischen Wellengleichung des Mehrteilchensystems aus einem klassischen Modell*, Zeitschrift für Physik A **136**, pp. 582–604, 1954. This work was taken up by E. Nelson, *Derivation of the Schrödinger equation from Newtonian mechanics*, Physical Review **150**, pp. 1079–1085, 1969, and in Edward Nelson, *Quantum Fluctuations*, Princeton University Press 1985, also downloadable at www.math.princeton.edu/~nelson/books.html, and the book Edward Nelson, *Stochastic Quantization*, Princeton University Press 1985. See also L. Fritsche & M. Haugk, *A new look at the derivation of the Schrödinger equation from Newtonian mechanics*, Annalen der Physik **12**, pp. 371–402, 2003. A summary of Nelson's approach is also given in F. Markopoulou & L. Smolin, *Quantum theory from quantum gravity*, Physical Review D **70**, p. 124029, 2004, preprint at arxiv.org/abs/gr-qc/0311059. See also the important criticism by T. C. Wallstrom, *Inequivalence between the Schrödinger equation and the Madelung hydrodynamic equation*, Physical Review A **49**, pp. 1613–1617, 1994, and T. C. Wallstrom, *The stochastic mechanics of the Pauli equation*, Transactions of the American Mathematical Society **318**, pp. 749–762, 1990. A proposed answer is L. Smolin, *Could quantum mechanics be an approximation to another theory?*, preprint at arxiv.org/quant-ph/abs/0609109. See also S. K. Srinivasan & E. C. G. Sudarshan, *A direct derivation of the Dirac equation via quaternion measures*, Journal of Physics A **29**, pp. 5181–5186, 1996. Cited on page 220.

177 Julian Schwinger, *Quantum Mechanics – Symbolism of Atomic Measurements*, Springer, 2001. Cited on page 222.

178 H. Nikolić, *How (not) to teach Lorentz covariance of the Dirac equation*, European Journal of Physics **35**, p. 035003, 2014, preprint at arxiv.org/abs/1309.7070. Cited on page 222.

179 For such an attempt, see the proposal by M. Rainer, *Resolution of simple singularities yielding particle symmetries in space-time*, Journal of Mathematical Physics **35**, pp. 646–655, 1994. Cited on page 225.

180 C. Schiller, *Deducing the three gauge interactions from the three Reidemeister moves*, preprint at arxiv.org/abs/0905.3905. Cited on pages 225 and 228.

181 G. T. Horowitz & J. Polchinski, *Gauge/gravity duality*, preprint at arxiv.org/abs/gr-qc/0602037. Note also the statement in the introduction that a graviton might be a composite of two spin-1 bosons, which is somewhat reproduced by the strand model of the graviton. A more concrete approach to gauge–gravity duality is made by M. van Raamsdonk, *Building up spacetime with quantum entanglement*, preprint at arxiv.org/1005.3035. This approach to gauge–gravity duality is close to that of the strand model. No citations.

182 K. Reidemeister, *Elementare Begründung der Knotentheorie*, Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg **5**, pp. 24–32, 1926. Cited on pages 228 and 276.

183 L. Burns, *Maxwell's Equations are Universal for* Locally *Conserved Quantities*, Advances in Applied Clifford Algebras **29**, p. 62, 2019. See also J. A. Heras, *Can Maxwell's equations be obtained from the continuity equation?*, American Journal of Physics 75, pp. 652–657, 2007. Cited on pages 229, 235, and 243.

184 For an attempt to reconcile braided particle models and SU(5) GUT, see D. Cartin, *Braids as a representation space of SU(5)*, preprint at arxiv.org/pdf/1506.08067. Cited on page 246.

185 Sheldon Glashow, confirmed this to the author in an email; Richard Feynman, makes the point in James Gleick, *Genius: The Life and Science of Richard Feynman*, Vintage Books, 1991, page 288 and also in Robert Crease & Charles Mann, *The Second Creation: Makers of the Revolution in Twentieth-Century Physics*, Macmillan Publishing, page 418; Martin Veltman, writes this in his Nobel Prize Lecture, available on www.nobel.org. Cited on page 257.

186 For some of the background on this topic, see F. Wilczek & A. Zee *Appearance of gauge structures in simple dynamical systems*, Physical Review Letters 52, pp. 2111–2114, 1984, A. Shapere & F. Wilczek, *Self-propulsion at low Reynolds number*, Physical Review Letters 58, pp. 2051–2054, 1987, and A. Shapere & F. Wilczek, *Gauge kinematics of deformable bodies*, American Journal of Physics 57, pp. 514–518, 1989. Cited on page 274.

187 R. Britto, F. Cachazo, B. Feng & E. Witten, *Direct proof of tree-level recursion relation in Yang–Mills theory*, preprint at arxiv.org/abs/hep-th/0501052. Cited on page 275.

188 D. V. Ahluwalia-Khalilova, *Operational indistinguishability of double special relativity from special relativity*, Classical and Quantum Gravity 22, pp. 1433–1450, 2005, preprint at arxiv.org/abs/gr-qc/0212128; see also N. Jafari & A. Shariati, *Doubly special relativity: a new relativity or not?*, preprint at arxiv.org/abs/gr-qc/0602075. Cited on page 282.

189 E. Verlinde, *On the origin of gravity and the laws of Newton*, preprint at arxiv.org/abs/1001.0785. Cited on page 283.

190 G.-L. Lesage, *Lucrèce Newtonien*, Nouveaux mémoires de l'Académie Royale des Sciences et Belles Lettres pp. 404–431, 1747, or www3.bbaw.de/bibliothek/digital/struktur/03-nouv/1782/jpg-0600/00000495.htm. See also en.wikipedia.org/wiki/Le_Sage's_theory_of_gravitation. In fact, the first to propose the idea of gravitation as a result of small particles pushing masses was Nicolas Fatio de Duillier in 1688. Cited on page 285.

191 G. 't Hooft, *Dimensional reduction in quantum gravity*, preprint at arxiv.org/abs/gr-qc/9310026. Many of the ideas of this paper become easier to understand and to argue when the strand model is used. Cited on page 290.

192 S. Carlip, *Logarithmic corrections to black hole entropy from the Cardy formula*, Classical and Quantum Gravity 17, pp. 4175–4186, 2000, preprint at arxiv.org/abs/gr-qc/0005017. Cited on page 291.

193 D. N. Page, *The Bekenstein Bound*, preprint at arxiv.org/abs/1804.10623. Cited on page 293.

194 On the limit for angular momentum of black holes, see Ref. 34. Cited on page 293.

195 F. Tamburini, C. Cuofano, M. Della Valle & R. Gilmozzi, *No quantum gravity signature from the farthest quasars*, preprint at arxiv.org/abs/1108.6005. Cited on page 298.

196 B.P. Abbott & al., (LIGO Scientific Collaboration and Virgo Collaboration) *Observation of gravitational waves from a binary black hole merger*, Physical Review Letters 116, p. 061102, 2016, also available for free download at journals.aps.org/prl/pdf/10.1103/PhysRevLett.116.061102. See also the website www.ligo.caltech.edu. More about this discovery and its implications is told in volume II of the Motion Mountain series. Cited on page 299.

197 On torsion, see the excellent review by R. T. Hammond, *New fields in general relativity*, Contemporary Physics 36, pp. 103–114, 1995. Cited on page 301.

198 H. Kleinert, & J. Zaanen, *World nematic crystal model of gravity explaining the absence of torsion*, Physics Letters A 324, pp. 361–365, 2004. Cited on page 301.

199 The analogy between the situation around line defects and general relativity is explained in Ekkehart Kröner, *Kontinuumstheorie der Versetzungen und Eigenspannungen*, Springer, 1958, These ideas have been taken up and pursued by J. D. Eshelby, B. A. Bilby, and many others after them. Cited on page 301.

200 Loop quantum gravity is a vast research field. The complete literature is available at arxiv.org/archive/gr-qc. Cited on page 301.

201 G. 't Hooft, *Crystalline Gravity*, International Journal of Modern Physics A **24**, pp. 3243–3255, 2009, and also G. 't Hooft, *A locally finite model of gravity*, preprint at arxiv.org/abs/0804.0328. Cited on page 301.

202 L. Susskind, *New concepts for old black holes*, preprint at arxiv.org/abs/1311.3335, and also reference Ref. 147. Cited on page 302.

203 M. Botta Cantcheff, *Spacetime geometry as statistic ensemble of strings*, preprint at arxiv.org/abs/1105.3658. Cited on page 302.

204 N. Arkani-Hamed, L. Motl, A. Nicolis & C. Vafa, *The string landscape, black holes and gravity as the weakest force*, preprint at arxiv.org/abs/hep-th/0601001. The paper contradicts the strand model in multiple ways. Cited on page 303.

205 M. van Raamsdonk, *Comments on quantum gravity and entanglement*, preprint at arxiv.org/abs/0907.2939. Cited on page 303.

206 W. H. Zurek & K. S. Thorne, *Statistical mechanical origin of the entropy of a rotating, charged black hole*, Physical Review Letters **54**, pp. 2171–2175, 1985. Cited on page 303.

207 M. Shaposhnikov & C. Wetterich, *Asymptotic safety of gravity and the Higgs boson mass*, preprint at arxiv.org/abs/0912.0208. Cited on page 304.

208 M. M. Anber & J. F. Donoghue, *On the running of the gravitational constant*, preprint at arxiv.org/abs/1111.2875. Cited on page 304.

209 The 2016 data about modified Newtonian dynamics is found in S. McGaugh, F. Lelli & J. Schombert, *The radial acceleration relation in rotationally supported galaxies*, preprint at arxiv.org/abs/1609.05917, and in F. Lelli, S. S. McGaugh, J. M. Schombert & M. S. Pawlowski, *One law to rule them all: the radial acceleration relation of galaxies*, preprint at arxiv.org/abs/1610.08981. Cited on page 305.

210 C. H. Lineweaver & T. M. Davis, *Misconceptions about the big bang*, Scientific American pp. 36–45, March 2005. Cited on page 307.

211 Supernova Search Team Collaboration, A.G. Riess & al., *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, Astronomical Journal **116**, pp. 1009–1038, 1998, preprint at arxiv.org/abs/astro-ph/9805201. Cited on page 307.

212 Stephen Hawking & Roger Penrose, *The Nature of Space and Time*, Princeton University Press, 1996. Cited on page 308.

213 C. Balázs & I. Szapudi, *Naturalness of the vacuum energy in holographic theories*, preprint at arxiv.org/abs/hep-th/0603133. See also C. Bambi & F. R. Urban, *Natural extension of the generalised uncertainty principle*, preprint at arxiv.org/abs/0709.1965. The same point is made by D. A. Easson, P. H. Frampton & G. F. Smoot, *Entropic accelerating universe*, preprint at arxiv.org/abs/1002.4278. No citations.

214 W. Fischler & L. Susskind, *Holography and Cosmology*, preprint at arxiv.org/abs/hep-th/9806039. No citations.

215 For a review of recent cosmological data, see D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, G. Hinshaw, N. Jarosik, E. Komatsu,

L. Page, H. V. Peiris, L. Verde, C. Barnes, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack & E. L. Wright, *Wilkinson Microwave Anisotropy Probe (WMAP) three year results: implications for cosmology*, preprint at arxiv.org/abs/astro-ph/0603449. Cited on page 309.

216 There is a large body of literature that has explored a time-varying cosmological constant, especially in relation to holography. An example with many references is L. Xu, J. Lu & W. Li, *Time variable cosmological constants from the age of the universe*, preprint at arxiv.org/abs/0905.4773. No citations.

217 D. Wiltshire, *Gravitational energy and cosmic acceleration*, preprint at arxiv.org/abs/0712.3982 and D. Wiltshire, *Dark energy without dark energy*, preprint at arxiv.org/abs/0712.3984. A newer and well-argued paper is D. L. Wiltshire, *Cosmic structure, averaging and dark energy*, preprint at arxiv.org/abs/1311.3787. See also T. Buchert, A. A. Coley, H. Kleinert, B. F. Roukema & D. L. Wiltshire, *Observational challenges for the standard FLRW model*, preprint at arxiv.org/abs/1512.03313. No citations.

218 The attribution to Voltaire could not be confirmed. Cited on page 313.

219 V. Crede & C. A. Meyer, *The experimental status of glueballs*, Progress in Particle and Nuclear Physics **63**, pp. 74–116, 2009. Cited on page 344.

220 E. Klempt & A. Zaitsev, *Glueballs, hybrids, multiquarks. Experimental facts versus QCD inspired concepts*, Physics Reports **454**, 2007, preprint at arxiv.org/abs/0708.4016. Cited on pages 342 and 344.

221 R. V. Buniy & T. W. Kephart, *A model of glueballs*, preprint at arxiv.org/pdf/hep-ph/0209339; R. V. Buniy & T. W. Kephart, *Universal energy spectrum of tight knots and links in physics*, preprint at arxiv.org/pdf/hep-ph/0408025; R. V. Buniy & T. W. Kephart, *Glueballs and the universal energy spectrum of tight knots and links*, preprint at arxiv.org/pdf/hep-ph/0408027. See also J. P. Ralston, *The Bohr atom of glueballs*, preprint at arxiv.org/pdf/hep-ph/0301089. Cited on page 344.

222 A. J. Niemi, *Are glueballs knotted closed strings?*, pp. 127–129, in H. Suganuma, N. Ishii, M. Oka, H. Enyo, T. Hatsuda, T. Kunihiro & K. Yazaki editors, *Color confinement and hadrons in quantum chromodynamics*, World Scientific, 2003, preprint at arxiv.org/pdf/hep-th/0312133. See also Y. M. Cho, B. S. Park & P. M. Zhang, *New interpretation of Skyrme theory*, preprint at arxiv.org/pdf/hep-th/0404181; K. Kondo, A. Ono, A. Shibata, T. Shinohara & T. Murakami, *Glueball mass from quantized knot solitons and gauge-invariant gluon mass*, Jornal of Physics A **39**, pp. 13767–13782, 2006, preprint at arxiv.org/abs/hep-th/0604006. Cited on page 344.

223 See the one million dollar prize described at www.claymath.org/millennium/Yang-Mills_Theory. Cited on page 345.

224 For a clear review on the topic and the planned experiments, see E. Fiorini, *Measurement of neutrino mass in double beta decay*, Europhysics News **38**, pp. 30–34, 2007, downloadable at www.europhysicsnews.org. Cited on page 327.

225 For example, see the detailed discussion of neutrino properties at pdg.web.cern.ch or, in print, in Ref. 233. Cited on page 327.

226 For a possible third approach, see A. F. Nicholson & D. C. Kennedy, *Electroweak theory without Higgs bosons*, International Journal of Modern Physics A **15**, pp. 1497–1519, 2000, preprint at arxiv.org/abs/hep-ph/9706471. Cited on page 329.

227 M. Veltman, *The Higgs system*, lecture slides at www.nikhef.nl/pub/theory/academiclectures/Higgs.pdf. See also his CERN Yellow Report 97-05, *Reflections on the*

Higgs system, 1997, and the paper H. Veltman & M. Veltman, *On the possibility of resonances in longitudinally polarized vector boson scattering*, Acta Physics Polonica B **22**, pp. 669–695, 1991. Cited on page 329.

228 J. W. Moffat & V. T. Thot, *A finite electroweak model without a Higgs particle*, preprint at arxiv.org/abs/0812.1991. The ideas go back to D. Evens, J. W. Moffat, G. Kleppe & R. P. Woodard, *Nonlocal regularizations of gauge theories*, Physical Review D **43**, pp. 499–519, 1991. For more details on how to introduce non-locality while maintaining current conservation and unitarity, see G. Kleppe & R. P. Woodard, *Non-local Yang-Mills*, Nuclear Physics B **388**, pp. 81–112, 1992, preprint at arxiv.org/abs/hep-th/9203016. For a different approach that postulates no specific origin for the W and Z masses, see J. W. Moffat, *Ultraviolet complete electroweak model without a Higgs particle*, preprint at arxiv.org/abs/1006.1859. Cited on page 329.

229 H. B. Nielsen & P. Olesen, *A vortex line model for dual strings*, Nuclear Physics B **61**, pp. 45–61, 1973. Cited on pages 333 and 386.

230 B. Andersson, G. Gustafson, G. Ingelman & T. Sjöstrand, *Parton fragmentation and string dynamics*, Physics Reports **97**, pp. 31–145, 1983. Cited on page 333.

231 C. B. Thorn, *Subcritical string and large N QCD*, preprint at arxiv.org/abs/0809.1085. Cited on page 333.

232 A. J. Buchmann & E. M. Henley, *Intrinsic quadrupole moment of the nucleon*, Physical Review C **63**, p. 015202, 2000. Alfons Buchmann also predicts that the quadrupole moment of the other, strange $J = 1/2$ octet baryons is positive, and predicts a prolate structure for all of them (private communication). For the decuplet baryons, with $J = 3/2$, the quadrupole moment can often be measured spectroscopically, and is always negative. The four Δ baryons are thus predicted to have a negative intrinsic quadrupole moment and thus an oblate shape. This explained in A. J. Buchmann & E. M. Henley, *Quadrupole moments of baryons*, Physical Review D **65**, p. 073017, 2002. For recent updates, see A. J. Buchmann, *Charge form factors and nucleon shape*, pp. 110–125, in C. N. Papanicolas & Aron Bernstein editors, *Shape of Hadrons Workshop Conference*, Athens, Greece, 27-29 April 2006, AIP Conference Proceedings 904. Cited on pages 336 and 340.

233 C. Patrignani & al., (Particle Data Group), Chinese Physics C **40**, p. 100001, 2016, or pdg.web.cern.ch. Cited on pages 337, 338, 339, 361, 363, 376, 377, 378, 379, 464, and 466.

234 A review on Regge trajectories and Chew-Frautschi plots is W. Drechsler, *Das Regge-Pol-Modell*, Naturwissenschaften **59**, pp. 325–336, 1972. See also the short lecture on courses.washington.edu/phys55x/Physics557_lec11.htm. Cited on page 337.

235 Kurt Gottfried & Victor F. Weisskopf, *Concepts of Particle Physics*, Clarendon Press, Oxford, 1984. Cited on page 338.

236 G. 't Hooft, G. Isidori, L. Maiani, A. D. Polosa & V. Riquer, *A theory of scalar mesons*, Physics Letters B **662**, pp. 424–430, 2008, preprint at arxiv.org/abs/0801.2288. However, other researchers, such as arxiv.org/abs/1404.5673, argue against the tetraquark interpretation. The issue is not closed. Cited on page 342.

237 M. Karliner, *Doubly heavy tetraquarks and baryons*, preprint at arxiv.org/abs/1401.4058. Cited on page 346.

238 J. Viro & O. Viro, *Configurations of skew lines*, Leningrad Mathematical Journal **1**, pp. 1027–1050, 1990, and updated preprint at arxiv.org/abs/math.GT/0611374. Cited on page 347.

239 W. Thomson, *On vortex motion*, Transactions of the Royal Society in Edinburgh pp. 217–260, 1868. This famous paper stimulated much work on knot theory. Cited on page 347.

240 H. Jehle, *Flux quantization and particle physics*, Physical Review D **6**, pp. 441–457, 1972, and H. Jehle, *Flux quantization and fractional charge of quarks*, Physical Review D **6**, pp. 2147–2177, 1975. Cited on page 347.

241 T. R. Mongan, *A holographic charged preon model*, preprint at arxiv.org/abs/0801.3670. Cited on page 347.

242 The arguments can be found in A. H. Chamseddine, A. Connes & V. Mukhanov, *Geometry and the quantum: basics*, preprint at arxiv.org/abs/1411.0977 and in A. H. Chamseddine & A. Connes, *Why the standard model*, Journal of Geometry and Physics 58, pp. 38–47, 2008, preprint at arxiv.org/abs/0706.3688. Cited on page 348.

243 Jacob's rings are shown, for example, in the animation on www.prestidigitascience.fr/index.php?page=anneaux-de-jacob. They are already published in the book by Tom Tit, *La science amusante*, 1870, and the images were reprinted the popular science books by Edi Lammers, and, almost a century later on, even in the mathematics column and in one of the books by Martin Gardner. See also www.lhup.edu/~dsimanek/scenario/toytrick.htm. Cited on page 351.

244 R. Boughezal, J. B. Tausk & J. J. van der Bij, *Three-loop electroweak corrections to the W-boson mass and* $\sin^2\theta_{\text{eff}}$ *in the large Higgs mass limit*, Nuclear Physics B 725, pp. 3–14, 2005, preprint at arxiv.org/abs/hep-ph/0504092. Cited on page 361.

245 The topic of the g-factor of the W boson and of charged fermions is covered in the delightful paper by Barry R. Holstein, *How large is the "natural" magnetic moment?*, American Journal of Physics 74, pp. 1104–1111, 2006, preprint at arxiv.org/abs/hep-ph/0607187. Cited on page 363.

246 The calculations have been performed in August 2016 by Eric Rawdon. Cited on pages 361 and 363.

247 The calculations have been performed by Eric Rawdon and Maria Fisher. Cited on page 365.

248 The quark masses at Planck energy are due to a private communication by Xing Zhizhong and Zhou Shun. They are calculated following the method presented in *Quark mass hierarchy and flavor mixing puzzles*, preprint at arxiv.org/abs/1411.2713 and Zhizhong Xing, He Zhang & Shun Zhou, *Updated values of running quark and lepton masses*, preprint at arxiv.org/abs/0712.1419. Cited on page 365.

249 See H. Fritzsch, A.D. Özer, *A scaling law for quark masses*, preprint at arxiv.org/abs/hep-ph/0407308. Cited on page 366.

250 K. A. Meissner & H. Nicolai, *Neutrinos, axions and conformal symmetry*, preprint at arxiv.org/abs/0803.2814. Cited on pages 369 and 370.

251 M. Shaposhnikov, *Is there a new physics between electroweak and Planck scale?*, preprint at arxiv.org/abs/0708.3550. Cited on page 370.

252 Y. Diao, C. Ernst, A. Por & U. Ziegler, *The roplength of knots are almost linear in terms of their crossing numbers*, preprint at arxiv.org/abs/0912.3282. Cited on page 373.

253 H. Fritzsch & Z.-Z. Xing, *Lepton mass hierarchy and neutrino mixing*, preprint at arxiv.org/abs/hep-ph/0601104 Cited on page 378.

254 The effects of neutrino mixing, i.e., neutrino oscillations, were measured in numerous experiments from the 1960s onwards; most important were the experiments at Super-Kamiokande in Japan and at the Sudbury Neutrino Observatory in Canada. See Ref. 233. Cited on page 379.

255 M. Fukugita & T. Yanagida, *Baryogenesis without grand unification*, Physics Letters B **174**, pp. 45–47, 1986. Cited on page 380.

256 J. M. Cline, *Baryogenesis*, preprint at arxiv.org/abs/hep-ph/0609145 or the review by L. Canetti, M. Drewes & M. Shaposhnikov, *Matter and Antimatter in the Universe*, preprint at arxiv.org/abs/1204.4186. They explain the arguments that the standard model with its CKM-CP violation is not sufficient to explain baryogenesis. The opposite view, by the same authors, is found in L. Canetti, M. Drewes, T. Frossard & M. Shaposhnikov, *Dark matter, baryogenesis and neutrino oscillations from right handed neutrinos*, preprint at arxiv.org/abs/1208.4607; another opposing view is found in T. Brauner, *CP violation and electroweak baryogenesis in the Standard Model*, EPJ Web of Conferences **70**, p. 00078, 2014. This issue is not settled yet. Cited on page 380.

257 Several claims that the coupling constants changed with the age of the universe have appeared in the literature. The first claim was by J. K. Webb, V. V. Flambaum, C. W. Churchill, M. J. Drinkwater & J. D. Barrow, *A search for time variation of the fine structure constant*, Physical Review Letters **82**, pp. 884–887, 1999, preprint at arxiv.org/abs/astro-ph/9803165. Several subsequent claims have appeared. However, non of these claims has been confirmed by subsequent measurements. Nowadays, there is no data showing that the fine structure constant changes in time. Cited on page 384.

258 The poster on www.physicsoverflow.org referred to J. P. Lestone, *Physics based calculation of the fine structure constant*, preprint at arxiv.org/abs/physics/0703151. The preprint has never been published. Cited on page 389.

259 For a highly questionable, but still intriguing argument based on black hole thermodynamics that claims to deduce the limit $\alpha > \ln 3/48\pi \approx 1/137.26$, see S. Hod, *Gravitation, thermodynamics, and the fine-structure constant*, International Journal of Modern Physics D **19**, pp. 2319–2323, 2010. It might well be that similar or other arguments based on textbook physics will yield more convincing or even better limits in the future. Cited on page 389.

260 V. Arnold, *Topological Invariants of Plane Curves and Caustics*, American Mathematical Society, 1994. Cited on page 396.

261 See M. Pospelov & A. Ritz, *Electric dipole moments as probes of new physics*, preprint at arxiv.org/abs/hep-ph/0504231. Cited on page 396.

262 C. Schiller, *A conjecture on deducing general relativity and the standard model with its fundamental constants from rational tangles of strands*, Physics of Particles and Nuclei **50**, pp. 259–-299, 2019. Cited on pages 398 and 433.

263 D. Hilbert, *Über das Unendliche*, Mathematische Annalen **95**, pp. 161–190, 1925. Cited on page 418.

264 The *Book of Twenty-four Philosophers*, *c.* 1200, is attributed to the god Hermes Trismegistos, but was actually written in the middle ages. The text can be found in F. Hudry, ed., *Liber viginti quattuor philosophorum*, Turnholt, 1997, in the series *Corpus Christianorum, Continuatio Mediaevalis*, CXLIII a, tome III, part 1, of the Hermes Latinus edition project headed by P. Lucentini. There is a Spinozian cheat in the quote: instead of 'nature', the original says 'god'. The reason why this substitution is applicable is given above. Cited on page 423.

Page 441

265 As a disappointing example, see Gilles Deleuze, *Le Pli – Leibniz et le baroque*, Les Editions de Minuit, 1988. In this unintelligible, completely crazy book, the author pretends to investigate the implications of the idea that the *fold* (in French 'le pli') is the basic entity of matter and 'soul'. Cited on page 425.

266 Werner Heisenberg, *Der Teil und das Ganze*, Piper, 1969. The text shows well how boring the personal philosophy of an important physicist can be. Cited on page 426.

267 John Barrow wrote to the author saying that he might indeed have been the first to have used the T-shirt image, in his 1988 Gifford Lectures at Glasgow that were a precursor to his book John D. Barrow, *Theories of Everything: The Quest for Ultimate Explanation*, 1991. He added that one can never be sure, though. Cited on page 428.

268 René Descartes, *Discours de la méthode*, 1637. He used and discussed the sentence again in his *Méditations métaphysiques* 1641, and in his *Les principes de la philosophie* 1644. These books influenced many thinkers in the subsequent centuries. Cited on page 431.

269 D. D. Kelly, *Sleep and dreaming*, in *Principles of Neural Science*, Elsevier, New York, 1991. The paper summarises experiments made on numerous humans and shows that even during dreams, people's estimate of time duration corresponds to that measured by clocks. Cited on page 432.

270 Astrid Lindgren said this in 1977, in her speech at the fiftieth anniversary of Oetinger Verlag, her German publisher. The German original is: 'Alles was an Großem in der Welt geschah, vollzog sich zuerst in der Phantasie eines Menschen, und wie die Welt von morgen aussehen wird, hängt in großem Maß von der Einbildungskraft jener ab, die gerade jetzt lesen lernen.' The statement is found in Astrid Lindgren, *Deshalb brauchen Kinder Bücher*, Oetinger Almanach Nr. 15, p. 14, 1977. Cited on page 435.

CREDITS

Acknowledgments

This volume was first published in 2009. No other person helped developing or exploring the strand model, until in 2014, Sergei Fadeev suggested to rethink the strand models for the W and Z bosons. His suggestion triggered many improvements, including a much clearer relation between the three Reidemeister moves and the intermediate gauge bosons of the three gauge interactions. The results were first included in 2015, in edition 28. The calculation of the fine structure constant became more involved, but is still possible.

A few other people helped to achieve progress on specific issues or provided encouragement. In the first half of the text, stimulating discussions in the 1990s with Luca Bombelli helped structuring the chapter on the contradictions between general relativity and quantum theory, as well as the chapter on the difference between vacuum and matter. In the years up to 2005, stimulating discussions with Saverio Pascazio, Corrado Massa and especially Steven Carlip helped shaping the chapter on limit values.

The second half of the text, on the strand model, owes much to Louis Kauffman. The ideas found in his books and in his papers inspired the ideas of this text long before we met and exchanged mails. His papers – available on www2.math.uic.edu/~kauffman – and his books are all worth reading; among them is the fascinating paper *Knot Logic* and the wonderful book *Knots and Physics*, World Scientific, 1991. His ideas on knots, on quantum theory, on measurement, on particle physics, on set theory and on foundational issues convinced me that strands are a promising direction in the search for a complete theory of motion. The breadth of Louis Kauffman's interests and the depth of his passion are exemplary.

I thank Eric Rawdon and Maria Fisher for their ropelength calculations. I also thank Claus Ernst, Andrzej Stasiak, Ralf Metzler and Jason Cantarella for their input and the fruitful discussions we had.

Hans Aschauer, Roland Netz, Gerrit Bauer, Stephan Schiller, Richard Hoffmann, Axel Schenzle, Reinhard Winterhoff, Alden Mead, Franca Jones-Clerici, Damoon Saghian, Frank Sweetser, Franz Aichinger, Marcus Platzer, Miles Mutka, and a few people who want to remain anonymous provided valuable help. My parents, Isabella and Peter Schiller, strongly supported the project. I thank my mathematics and physics teachers in secondary school, Helmut Wunderling, for the fire he has nurtured inside me.

The typesetting and book design is due to the professional consulting of Ulrich Dirr. The typography was much improved with the help of Johannes Küster and his Minion Math font. The design of the book and its website also owe much to the suggestions and support of my wife Britta.

From 2007 to 2011, the electronic edition and distribution of the Motion Mountain text, and the research for this volume, was encouraged by Klaus Tschira and generously supported by the Klaus Tschira Foundation. I also thank the lawmakers and the taxpayers in Germany, who, in contrast to most other countries in the world, allow residents to use the local university libraries.

Film credits

The animations of the belt trick on page 178 are copyright and courtesy of Greg Egan; they can be found on his website www.gregegan.net/APPLETS/21/21.html. I am also grateful to Greg Egan for expanding his applet so as to show a second option out of the many possible ones for the belt trick.

The beautiful animation of the belt trick on page 179 and the wonderful and so-far unique animation of the fermion exchange on page 184 are copyright and courtesy of Antonio Martos. He made them for this text. They can be found at vimeo.com/62228139 and vimeo.com/62143283.

The animation of the rotating blue ball attached to a sheet on page 180 was made for this book and is part of the software found at www.ariwatch.com/VS/Algorithms/DiracStringTrick.htm.

The colourful animations of the belt trick with many tails on page 182 and page 182 are copyright and courtesy of Jason Hise at www.entropygames.net, and were made for this text and for the Wikimedia Commons website.

The film of the chain ring trick on page 350 is copyright and courtesy of Franz Aichinger.

The animations and the film of the falling chain ring trick were included into the pdf file with the help of a copy of the iShowU software sponsored by Neil Clayton from www.shinywhitebox.com.

Image credits

The photograph of the east side of the Langtang Lirung peak in the Nepalese Himalayas, shown on the front cover, is courtesy and copyright by Kevin Hite and found on his blog thegettingthere.com. The photograph of the Ultra Deep Field project on page 16 is courtesy of NASA. The drawing by Maurits Escher on page 63 is copyright by the M.C. Escher Heirs, c/o Cordon Art, Baarn, The Netherlands, who kindly gave permission for its use. The passport photograph of Andrei Sakharov on page 79 is copyright by himself. The photograph of an apheresis machine on page 181 is courtesy of Wikimedia. The drawing by Peter Battey-Pratt and Thomas Racey of the belt trick on page 217, taken from Ref. 172, is courtesy and copyright by Springer Verlag. The graph of the running coupling constants on page 382 is courtesy and copyright by Wim de Boer and taken from his home page at www-ekp.physik.uni-karlsruhe.de/~deboer. On page 427, the photograph of Cerro Torre is copyright and courtesy of Davide Brighenti, and found on Wikimedia; the photograph of the green hill on the same page is copyright and courtesy of Myriam70, and found on her site www.flickr.com/photos/myriam70. The photograph on the back cover, of a basilisk running over water, is courtesy and copyright by the Belgian group TERRA vzw and found on their website www.terravzw.org. All drawings are copyright by Christoph Schiller.

NAME INDEX

Z

SUBJECT INDEX

E

ELEMENTS

F

FINE-TUNING

G

GLUONIC

H

HUBBLE

L

Lagrangian

M

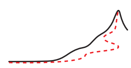
P

PLANCK

Q

R

R
REAL

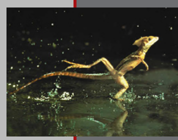


MOTION MOUNTAIN

The Adventure of Physics – Vol. VI

The Strand Model –

A Speculation on Unification



What is the origin of colours?
Which problems in physics are unsolved since the year 2000 and what might be their solution?
Why do motion and change exist?
What is the origin of the principle of least action?
What is the origin of gauge symmetries?
What is the most fantastic voyage possible?

Answering these and other questions, this book gives an entertaining and mind-twisting introduction to the search for the final theory of physics. The search leads to the *strand model*: Based on a simple principle, strands reproduce quantum theory, the standard model of particle physics and general relativity. Strands agree with all experimental data and allow estimating the fine structure constant, particle masses and all other constants of nature.

Christoph Schiller, PhD Université Libre de Bruxelles, is a physicist and physics popularizer. This entertaining book is for students, teachers and anybody interested in modern research about fundamental physics.

Pdf file available free of charge at
www.motionmountain.net/research.html