Christoph Schiller

MOTION MOUNTAIN

THE ADVENTURE OF PHYSICS – VOL.III

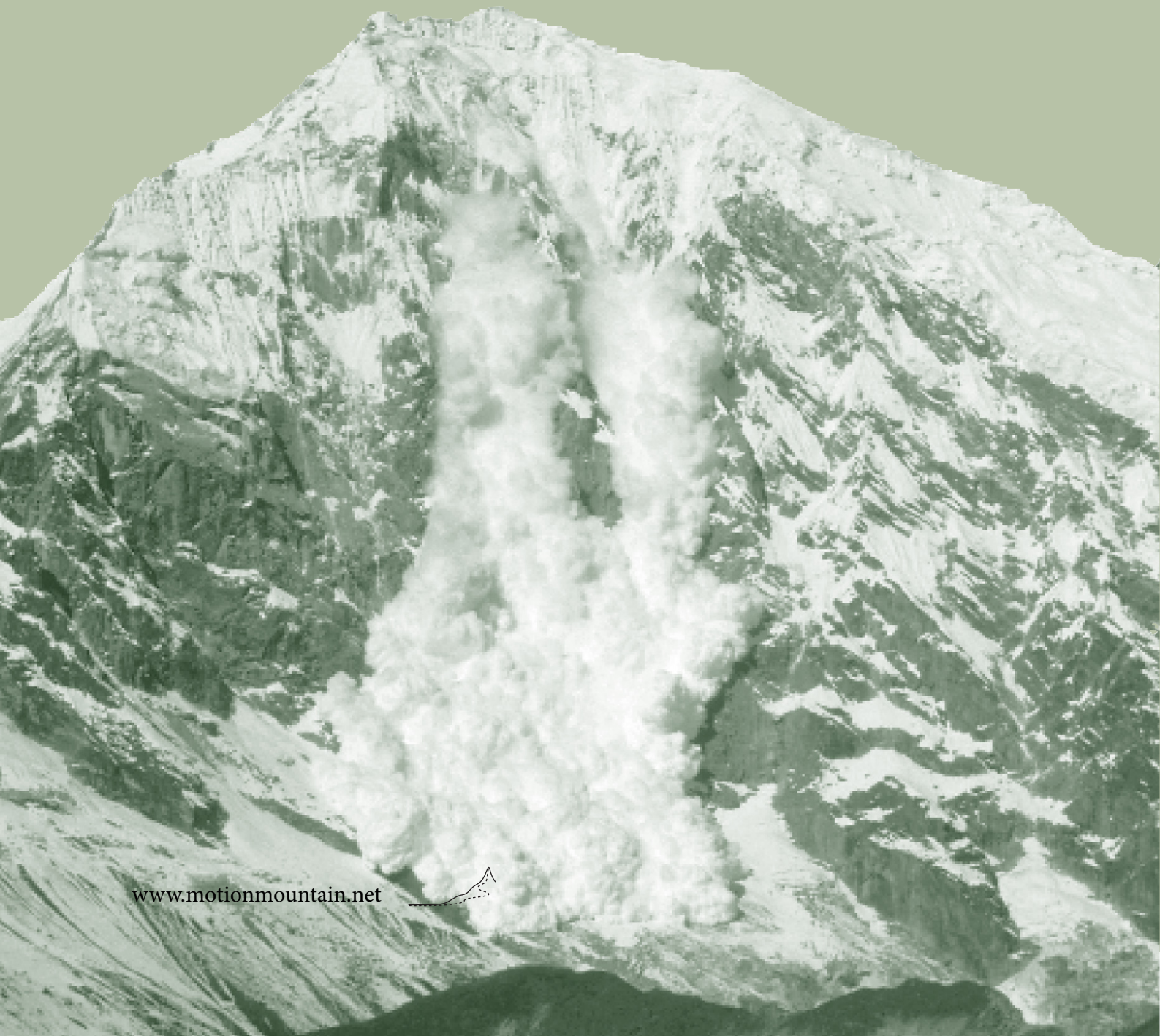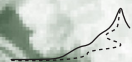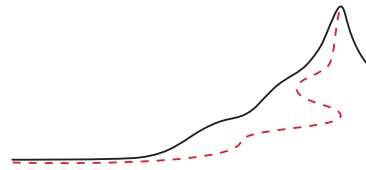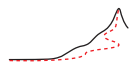### LIGHT, CHARGES AND BRAINS

www.motionmountain.net

Christoph Schiller

MOTION MOUNTAIN

The Adventure of Physics
Volume III

Light, Charges and Brains

To Britta, Esther and Justus Aaron

τῷ ἐμοὶ δαίμονι

Die Menschen stärken, die Sachen klären.

PREFACE

> “ Primum movere, deinde docere.*
>
> Antiquity ”

This book is written for anybody who is curious about nature and motion. Curiosity about how people, animals, things, images and space move leads to many adventures. This volume presents the adventures one encounters when exploring everything electric. The story ranges from the weighing of electric current to the use of magnetic fields to heal bone fractures and up to the understanding of the human brain.

In the structure of physics, shown in Figure 1, motion due to electricity is the most fascinating aspect of the starting point at the bottom. Indeed, almost everything around us is due to electric processes. The present introduction to electricity, magnetism, light and the brain is the third of a six-volume overview of physics that arose from a threefold aim that I have pursued since 1990: to present motion in a way that is simple, up to date and captivating.

In order to be *simple*, the text focuses on concepts, while keeping mathematics to the necessary minimum. Understanding the concepts of physics is given precedence over using formulae in calculations. The whole text is within the reach of an undergraduate.

In order to be *up to date*, the text is enriched by the many gems – both theoretical and empirical – that are scattered throughout the scientific literature.

In order to be *captivating*, the text tries to startle the reader as much as possible. Reading a book on general physics should be like going to a magic show. We watch, we are astonished, we do not believe our eyes, we think, and finally we understand the trick. When we look at nature, we often have the same experience. Indeed, every page presents at least one surprise or provocation for the reader to think about. Numerous interesting challenges are proposed.

The motto of the text, *die Menschen stärken, die Sachen klären*, a famous statement by Hartmut von Hentig on pedagogy, translates as: ‘To fortify people, to clarify things.’ Clarifying things requires courage, as changing habits of thought produces fear, often hidden by anger. But by overcoming our fears we grow in strength. And we experience intense and beautiful emotions. All great adventures in life allow this, and exploring motion is one of them.

Munich, 24 June 2011.

---

* ‘First move, then teach.’ In modern languages, the mentioned type of *moving* (the heart) is called *motivating*; both terms go back to the same Latin root.
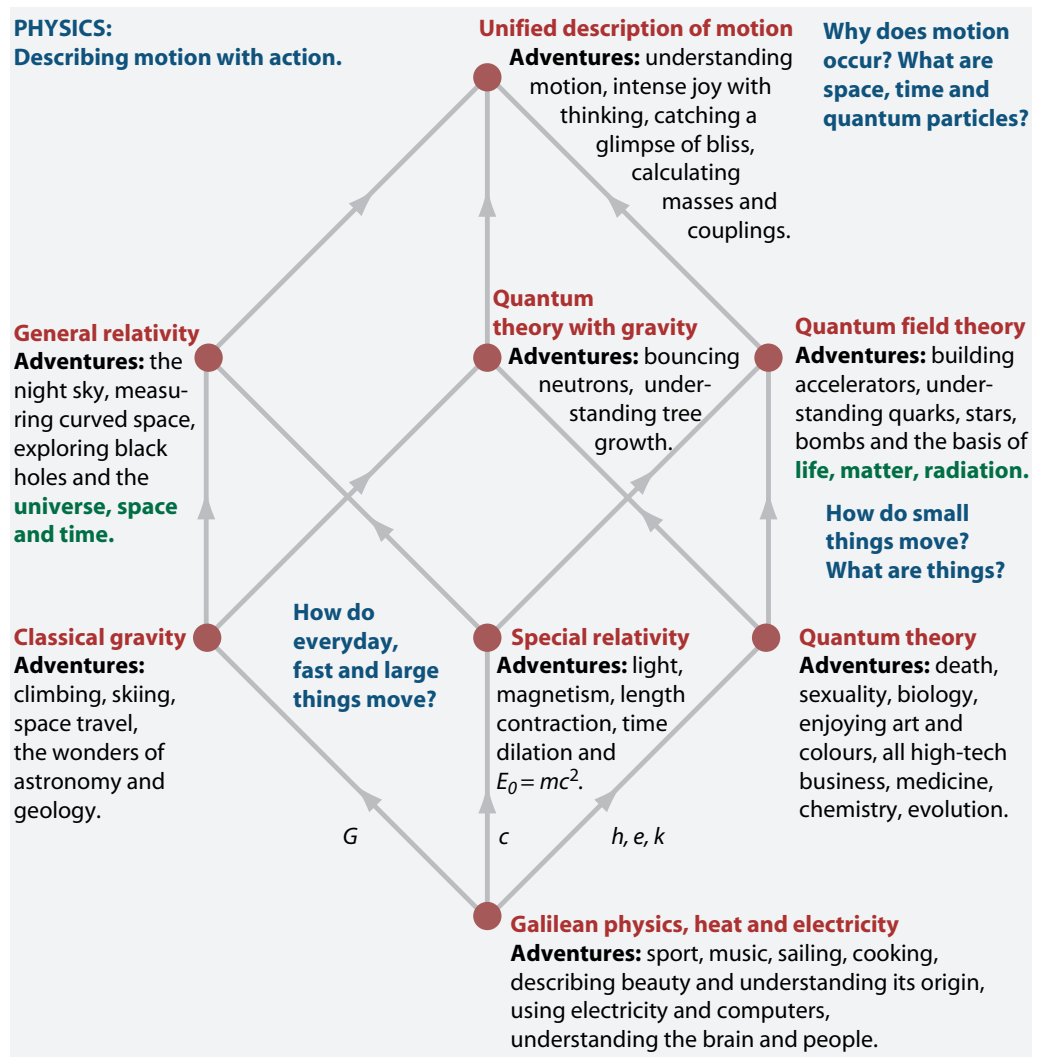
PHYSICS:
Describing motion with action.

Unified description of motion
Adventures: understanding motion, intense joy with thinking, catching a glimpse of bliss, calculating masses and couplings.

Why does motion occur? What are space, time and quantum particles?

General relativity
Adventures: the night sky, measuring curved space, exploring black holes and the **universe, space and time.**

Quantum theory with gravity
Adventures: bouncing neutrons, understanding tree growth.

Quantum field theory
Adventures: building accelerators, understanding quarks, stars, bombs and the basis of **life, matter, radiation.**

How do small things move? What are things?

Classical gravity
Adventures: climbing, skiing, space travel, the wonders of astronomy and geology.

How do everyday, fast and large things move?

Special relativity
Adventures: light, magnetism, length contraction, time dilation and $E_0 = mc^2$.

Quantum theory
Adventures: death, sexuality, biology, enjoying art and colours, all high-tech business, medicine, chemistry, evolution.

G c h, e, k

Galilean physics, heat and electricity
Adventures: sport, music, sailing, cooking, describing beauty and understanding its origin, using electricity and computers, understanding the brain and people.

FIGURE 1 A complete map of physics: the connections are defined by the speed of light c, the gravitational constant G, the Planck constant h, the Boltzmann constant k and the elementary charge e.

ADVICE FOR LEARNERS

In my experience as a teacher, there was one learning method that never failed to transform unsuccessful pupils into successful ones: if you read a book for study, summarize every section you read, *in your own words, aloud*. If you are unable to do so, read the section again. Repeat this until you can clearly summarize what you read in your own words, aloud. You can do this alone in a room, or with friends, or while walking. If you do this with everything you read, you will reduce your learning and reading time significantly. In addition, you will enjoy learning from good texts much more and hate bad texts much less. Masters of the method can use it even while listening to a lecture, in a low voice, thus avoiding to ever take notes.

Using this book

Text in green, as found in many marginal notes, marks a link that can be clicked in a pdf reader. Such green links are either bibliographic references, footnotes, cross references to other pages, challenge solutions, or pointers to websites.

Solutions and hints for *challenges* are given in the appendix. Challenges are classified as research level (r), difficult (d), standard student level (s) and easy (e). Challenges of type r, d or s for which no solution has yet been included in the book are marked (ny).
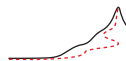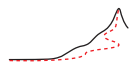
Feedback and support

This text is and will remain free to download from the internet. In exchange, I would be delighted to receive an email from you at fb@motionmountain.net, especially on the following issues:

Challenge 1 s   — What was unclear and should be improved?
— What story, topic, riddle, picture or movie did you miss?
— What should be corrected?

Alternatively, you can provide feedback online, on www.motionmountain.net/wiki. The feedback will be used to improve the next edition. On behalf of all readers, thank you in advance for your input. For a particularly useful contribution you will be mentioned – if you want – in the acknowledgements, receive a reward, or both.

Your donation to the charitable, tax-exempt non-profit organisation that produces, translates and publishes this book is welcome! For details, see the web page www.motionmountain.net/donation.html. If you want, your name will be included in the sponsor list. Thank you in advance for your help, in the name of all readers across the world. But above all, enjoy the reading!

Contents

Light, Charges and Brains

In our quest to learn how things move,
the experience of hiking and other motion
leads us to discover that images are produced by charges,
that charges move, accumulate and interact,
and that there is a smallest charge in nature.
We understand what love has to do with magnets and amber,
why the brain is such an interesting device,
and what distinguishes a good from a bad lie.
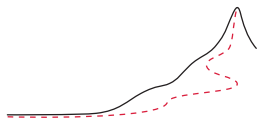
Chapter 1

LIQUID ELECTRICITY, INVISIBLE FIELDS AND MAXIMUM SPEED

What is *light*? The study of relativity left us completely in the dark, even though we had embarked in it precisely to find an answer to that question. True, we have learned how the motion of light compares with that of objects. We also learned that light is a moving entity that cannot be stopped, that light provides the speed limit for any type of energy, and that light is our measurement standard for speed. However, we haven't learned anything about the nature of light itself.

A second question is open: what is *contact*? We still do not know. The only thing we learned in our exploration of relativity was that truly mechanical interactions do not exist.
Indeed, all interactions are due to exchange of particles. But which ones?

The answer to the questions about the nature of light and contact emerges only from the study of those types of motion that are *not* related to gravitation. It turns out that the key to the answers is the understanding of the ways magicians levitate objects.

If we make a list of motors found in this world, we notice that gravitation hardly describes any type of motor. Neither the motion of sea waves, fire and earthquakes, nor that of a gentle breeze is caused by gravity. The same applies to the motion of muscles.*
Have you ever listened to your own heart beat with a stethoscope? (Or use, as many medical doctors do now, an mp3 player to record your heart beat.) Without having done so, you cannot claim to have experienced the mystery of motion. Your heart has about 3000 million beats in your lifetime. Then it stops.

It was one of the most astonishing discoveries of science that heart beats, sea waves and most other cases of everyday motion, as well as the nature of light itself, are connected to observations made thousands of years ago using two strange stones. These stones show that all those examples of motion that are called *mechanical* in everyday life are, without exception, of *electrical* origin.

In particular, the solidity, the softness and the impenetrability of matter are due to internal electricity; also the emission of light is an electrical process. As these aspects are part of everyday life, we will leave aside all complications due to gravity and curved space-time. The most productive way to study electrical motion is to start, as in the case of gravity, with those types of motion which are generated without any contact between the bodies involved.

---

* The photograph of a circular rainbow on page 14 was taken in 2006 from the Telstra Tower in Canberra (© Oat Vaiyaboon).

FIGURE 2 Objects surrounded by fields: amber (c. 1 cm), lodestone (c. 1 cm) and mobile phone (c. 10 cm) (© Philips).
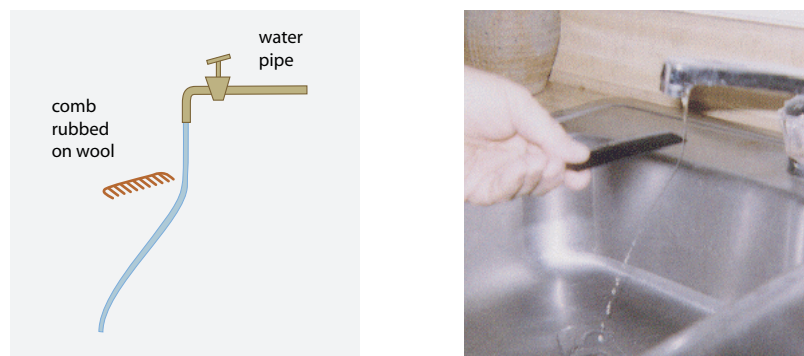
FIGURE 3 How to amaze kids, especially in dry weather (photo © Robert Fritzius).

Fields: amber, lodestone and mobile phones

The story of electricity starts with trees. Trees have a special relation to electricity. When a tree is cut, a viscous resin appears. With time it solidifies and, after millions of years, it forms *amber*. When amber is rubbed with a cat fur, it acquires the ability to attract small objects, such as saw dust or pieces of paper. This was already known to Thales of Miletus, one of the original seven sages, in the sixth century BCE. The same observation can be made with many other polymer combinations, for example with combs and hair, with soles of the shoe on carpets, and with a TV tube and dust. Children are always surprised by the effect, shown in Figure 3, that a comb rubbed on wool has on running tap water. The same effect can be produced with an air-filled rubber balloon rubbed on wool. Another interesting effect can be observed when a rubbed comb is put near a burning candle. (Can you imagine what happens?)

Challenge 3 ny

Another part of the story of electricity involves *lodestone*, an iron mineral found in certain caves around the world, e.g. in a region (still) called Magnesia in the Greek province of Thessalia, and in some regions in central Asia. When two stones of this mineral are put near each other, they attract or repel each other, depending on their relative orientation. In addition, lodestone attracts objects made of cobalt, nickel or iron.

Today we also find various small objects in nature with more sophisticated properties, such as the one shown on the right of Figure 2. Some of these objects allow you to talk

with far away friends, others unlock car doors, still others enable you to switch on a television.

All these observations show that in nature there are situations where bodies exert influence on others *at a distance*. The space surrounding a body exerting such an influence is said to contain a field. A *(physical) field* is thus an entity that manifests itself by accelerating other bodies in a given region of space. *A field is space that changes momenta*. If you prefer, *a field is space that exerts forces*. Or again, a field is space with some extra structure. Despite this extra structure, fields, like space, are invisible.

The field surrounding the mineral found in Magnesia is called a *magnetic field* and the stones are called *magnets*. The field around amber – called ἤλεκτρον in Greek, from a root meaning 'brilliant, shining' – is called an *electric field*. The name is due to a proposal by the famous English physician and part-time physicist William Gilbert (1544–1603). Objects surrounded by a permanent electric field are called *electrets*. Electrets are much less common than magnets; among others, they are used in certain loudspeaker systems.

The field around a mobile phone is called a *radio* field or, as we will see later, an *electromagnetic* field. In contrast to the previous fields, it oscillates over time. We will find out later that many other objects are surrounded by such fields, though these are often very weak. Objects that emit oscillating fields, such as mobile phones, are called radio transmitters or electromagnetic emitters.

Experiments show that fields have *no mass*. Without any material support, fields influence bodies over a distance. Fields are invisible. To make them imaginable, we just need to colour them. Some ways to colour electric fields are shown in Figure 4. These figures are the best way to *imagine* electric fields: they reproduce faithfully how the inventor of the field concept, Michael Faraday, imagined them.

For a long time, electric, magnetic and radio fields were rarely noticed in everyday life. Indeed, in the past, most countries had laws that did not allow producing such fields or building mobile phones or garage openers. Still today, laws severely restrict the properties of machines that use and produce such fields. The laws require that for any device that moves, produces sound, or creates moving pictures, fields need to remain *inside* the device. For this reason a magician moving an object on a table via a hidden magnet still surprises and entertains his audience. To feel the fascination of fields more strongly, a deeper look into a few experimental results is worthwhile.

HOW CAN ONE MAKE LIGHTNING?

Everybody has seen a lightning flash or has observed the effect it can have on striking a tree. Obviously lightning is a moving phenomenon. Photographs such as that of Figure 5 show that the tip of a lightning flash advance with an average speed of around 600 km/s. But *what* is moving? To find out, we have to find a way of making lightning for ourselves. In 1995, the car company General Motors accidentally rediscovered an old and simple method of achieving this.

Opel engineers had inadvertently built a spark generating mechanism into their cars; when filling the petrol tank, sparks were generated, which sometimes lead to the explosion of the fuel at the petrol station. They had to recall 2 million vehicles of its Opel brand.

What had the engineers done wrong? They had unwittingly copied the conditions for
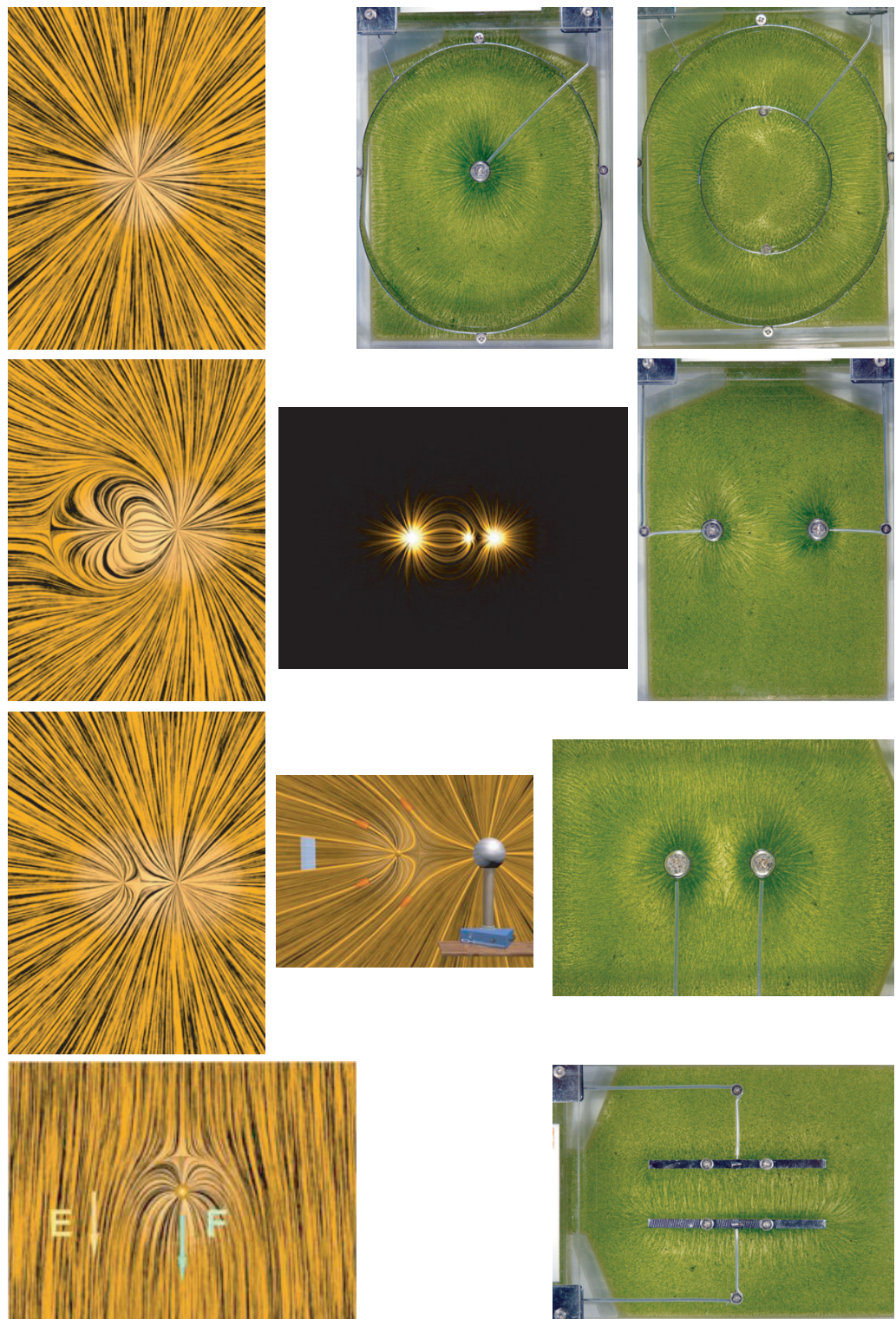
FIGURE 4 Visualizing what is invisible: a simple way to visualize electric fields as space with a structure, using computer graphics and using seeds in oil. Top: the field around a point or spherical charge; second row: two or three charges of different signs; third row: two charges of the same sign; bottom: a charge in an external field E, and the field between two plates. The charge will feel a force F directed along the so-called *electric field lines*; the density of the lines gives the intensity of the field and thus the strength of the force (© MIT, Eli Sidman, MIT).

FIGURE 5 Lightning: a picture taken with a moving camera, showing its multiple strokes (© Steven Horsburgh).

Ref. 4

a electrical device which anyone can build at home and which was originally invented by William Thomson:* the *Kelvin generator*. Repeating his experiment today, we would take two water taps, four empty bean or coffee cans, of which two have been opened at both sides, some nylon rope and some metal wire. Putting this all together as shown in Figure 6, and letting the water flow, we find a strange effect: large sparks periodically jump between the two copper wires at the point where they are nearest to each other, giving out loud bangs. Can you guess what condition for the flow has to be realized for

Challenge 4 s   this to work? And what did Opel do to repair the cars they recalled?

If we stop the water flowing just before the next spark is due, we find that both buckets are able to attract sawdust and pieces of paper. The generator thus does the same that rubbing amber does, just with more bang for the buck(et). Both buckets are surrounded by electric fields. The fields increase with time, until the spark jumps. Just after the spark, the buckets are (almost) without electric field. Obviously, the flow of water somehow collects something on each bucket; today we call this *electric charge*. Charge can flow in metals and, when the fields are high enough, through air. We also find that the two buckets are always surrounded by two different types of electric fields: bodies that are attracted by one bucket are repelled by the other.

All experiments confirm that there are *two* types of charges. The US politician and part-time physicist Benjamin Franklin (1706–1790) called the electricity created on a

---

* William Thomson (1824–1907), important Irish Unionist physicist and professor at Glasgow University. He worked on the determination of the age of the Earth, showing that it was much older than 6000 years, as several sects believed, but also (falsely) maintained that the Earth was much younger than geologists and Darwin (correctly) hat deduced. He strongly influenced the development of the theory of magnetism and electricity, the description of the aether and thermodynamics. He propagated the use of the term 'energy' as it is used today, instead of the confusing older terms. He was one of the last scientists to propagate mechanical analogies for the explanation of phenomena, and thus strongly opposed Maxwell's description of electromagnetism. It was mainly for this reason that he did not receive a Nobel Prize. He was also one of the minds behind the laying of the first transatlantic telegraphic cable. Victorian and religious to his bones, when he was knighted, he chose the name of a small brook near his home as his new name; thus he became Baron Kelvin of Largs. Therefore the unit of temperature obtained its name from a small Scottish river.
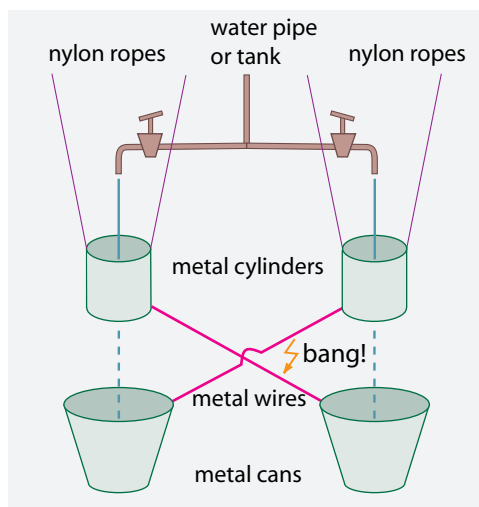
FIGURE 6 A simple Kelvin generator; the one on the right lights a fluorescent light bulb using dripping water (photograph © Harald Chmela).
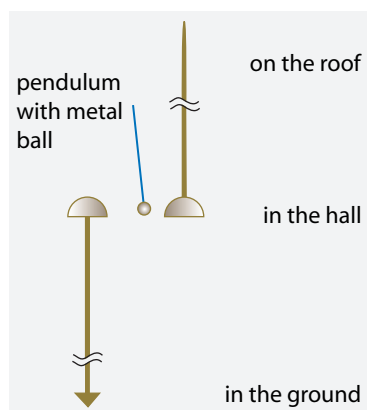
FIGURE 7 Franklin's personal lightning rod, a copy of Gordon's electric chime, one of the many experiments that shows visually that charge can flow.

glass rod rubbed with a dry cloth *positive*, and that on a piece of amber *negative*. (Previously, the two types of charges were called 'vitreous' and 'resinous'.) Bodies with charges of the same sign repel each other, bodies with opposite charges attract each other; charges of opposite sign flowing together cancel each other out.*

All flows take time. How fast is electricity? A simple way to measure the speed of electricity is to produce a small spark at one end of a long wire, and to observe how long it takes until the spark appears at the other end of the wire. In practice, the two sparks are almost simultaneous; the speed one measures is much higher than everything else we observe in our environment. How can we measure the time nevertheless? And why did different researchers get very different speed values in this experiment?

Challenge 5 s

---

* In fact, there are many other ways to produces sparks or even *arcs*, i.e., sustained sparks; there is even a complete subculture of people who do this as a hobby at home. Those who have a larger budget do it professionally, in particle accelerators. See the www.kronjaeger.com/hv website.

FIGURE 8 A simple set-up to confirm electric charge conservation: if rubbed fur is moved from the first pot to the second, the charge taken away from the first pot is transferred to the second, as shown by the two electrometers (© Wolfgang Rueckner).

Sparks, electric arcs and lightning are similar. Of course, one has to check whether natural lightning is actually electrical in origin. In 1752, experiments performed in France, following a suggestion by Benjamin Franklin, published in London in 1751, showed that one can indeed draw electricity from a thunderstorm via a long rod.* Thunderstorm clouds are surrounded by electric fields. These French experiments made Franklin famous worldwide; they were also the start of the use of lightning rods all over the world. Later, Franklin had a lightning rod built through his own house, but of a somewhat unusual type, as shown in Figure 7. This device, invented by Andrew Gordon, is called an *electric chime*. Can you guess what it did in his hall during bad weather, all parts being made of metal, and why? (Do not repeat this experiment; any device attached to a lightning rod can kill.)

Ref. 5

Challenge 6 s

In summary, electric fields start at bodies, provided they are charged. Charging can be achieved by rubbing and other processes. There are two types of charge, negative and positive. Charge can flow: it is then called an electric *current*. The worst conductors of current are polymers; they are called *insulators* or *dielectrics*. A charge put on an insulator remains at the place where it was put. In contrast, metals are good conductors; a charge placed on a conductor spreads all over its surface. The best conductors are silver and copper. This is the reason that at present, after two hundred years of use of electricity, the highest concentration of copper in the world is below the surface of Manhattan.

Electric charge

If all experiments with charge can be explained by calling the two charges positive and negative, the implication is that some bodies have more, and some less charge than an uncharged, *neutral* body. Electricity thus only flows when two differently charged bodies are brought into contact. Now, if charge can flow and accumulate, we must be able to somehow measure its amount. Obviously, the *amount* of electric charge on a body, usu-

---

* The details of how lightning is generated and how it propagates are still a topic of research. An introduction is given on page 147.

FIGURE 9 Various electrometers: a self-made electrometer based on a jam pot, an ancient (opened) high precision Dolezalek electrometer, the *Ampullae of Lorenzini* of a shark, and a modern digital electrometer (© Harald Chmela, Klaus Jost at www.jostimages.com, Advantest).

ally abbreviated q, must be defined via the influence the body, say a piece of sawdust, feels when subjected to a field. Charge is thus defined by comparing it to a standard reference charge. For a charged body of mass m accelerated in a field, its charge q is determined by the relation

$$\frac{q}{q_{\text{ref}}} = \frac{\mathrm{d}\mathbf{p}/\mathrm{d}t}{\mathrm{d}\mathbf{p}_{\text{ref}}/\mathrm{d}t} \,, \tag{1}$$

i.e., by comparing its momentum change with the momentum change of the reference frame. Charge thus determines the motion of bodies in electric fields in the same way that mass determines the motion of bodies in gravitational fields. Charge is therefore the second intrinsic property of bodies that we discover in our walk.

In practice, electric charge is measured with *electrometers*. A few such devices are shown in Figure 9. The main experimental properties of electric charge that are discovered when experimenting with electrometers are listed in Table 1. In all details, charge behaves like a flowing substance; charge behaves like a fluid.

Nowadays the unit of charge, the *coulomb*, is defined through a standard flow through metal wires, as explained in Appendix A. This is possible because all experiments show

TABLE 1 Properties of *classical* electric charge: a scalar density.

ELECTRIC CHARGES	PHYSICAL PROPERTY	MATHEMATICAL NAME	DEFINITION
Can be distinguished	distinguishability	element of set	Page 195
Can be ordered	sequence	order	Vol. IV, page 190
Can be compared	measurability	metricity	Vol. V, page 276
Can change gradually	continuity	completeness	Vol. V, page 284
Can be added	accumulability	additivity	Vol. I, page 73
Can be separated	separability	positive or negative	
Do not change	conservation	invariance	$q = \text{const}$

TABLE 2 Values of electrical charge observed in nature.

OBSERVATION	CHARGE
Smallest measured non-vanishing charge	$1.6 \cdot 10^{-19}$ C
Charge per bit in computer memory	down to 10^{-15} C
Charge in small capacitor	10^{-7} C
Charge flow in average lightning stroke	1 C to 100 C
Charge stored in a fully charged car battery	0.2 MC
Charge of planet Earth	1 MC
Charge separated by modern power station in one year	$3 \cdot 10^{11}$ C
Total charge of positive (or negative) sign observed in universe	$10^{60\pm1}$ C
Total charge observed in universe	0 C

that charge is *conserved*, that it *flows*, and thus that it can *accumulate*. In other words, if the electric charge of a physical system changes, the reason always is that charge is flowing into or out of the system. This can be checked easily with two metal pots connected

Ref. 6   to two electrometers, as shown in Figure 8. Charge thus behaves like a fluid substance. Therefore we are forced to use for its description a scalar quantity q, which can take positive, vanishing, or negative values.

Describing charge as a scalar quantity reproduces the behaviour electrical charge in all everyday situations. However, as in the case of all previously encountered classical concepts, some of the experimental results for electrical charge in everyday situations will turn out to be only approximate. More precise experiments will require a revision of the idea of continuous change of charge value. However, the main observation remains: no counter-example to charge conservation has as yet been observed.

Objects without electric charge are called *neutral*. A charged object that is brought near a neutral body polarizes it. *Electrical polarization* is the separation of the positive and negative charges in a body. For this reason, all neutral objects, such as hair, are attracted to a charged body, such as a rubbed comb. Generally, both insulators and conductors can be polarized; polarization occurs for single molecule up to whole stars.

Electric field strength

Charges produce attraction and repulsion on other charges. Equivalently, charges change momenta; charges exert forces on other charges. This happens over large distances. Experiments that explore energy and momentum conservation show that the best description of these interactions is as told so far: a charge produces a field, the field then acts on a second charge.

Experiments show that the *electric field* forms lines in space. As a consequence, the electric field behaves like a small arrow fixed at each point x in space. Electric fields are described by a direction and a magnitude. The local direction of the field is given by the local direction of the field line – the tangent of the field line. The local magnitude of the field is given by the local density of the field lines. The direction and the magnitude do not depend on the observer. In short, the electric field $E(x)$ is a *vector* field. Experiments show that it is best defined by the relation

$$q E(x) = \frac{\mathrm{d} p(x)}{\mathrm{d} t} \tag{2}$$

taken at every point in space x. The definition of the electric field is thus based on how it *moves* charges. In general, the electric field is a vector

$$E(x) = (E_x, E_y, E_z) \tag{3}$$

and is measured in multiples of the unit N/C or V/m.

The definition of the electric field assumes that the test charge q is so small that it does not disturb the field E. We sweep this issue under the carpet for the time being. This is a drastic move: we ignore quantum theory and all quantum effects in this way; we come back to it below.

The definition of the electric field also assumes that space-time is flat, and it ignores all issues due to space-time curvature.

By the way, does the definition of electric field just given assume a charge speed that is much less than that of light?

To describe the motion due to electricity completely, we need a relation explaining how charges *produce* electric fields. This relation was established with precision (but not for the first time) during the French Revolution by Charles-Augustin de Coulomb, on his private estate.* He found that around any small-sized or any spherical charge Q *at rest* there is an electric field. At a position r, the electric field E is given by

$$E(r) = \frac{1}{4\pi\varepsilon_0} \frac{Q}{r^2} \frac{r}{r} \quad \text{where} \quad \frac{1}{4\pi\varepsilon_0} = 9.0\,\mathrm{GV\,m/C} . \tag{4}$$

Later we will extend the relation for a charge in motion. The bizarre proportionality constant, built around the so-called *permittivity of free space* ε_0, is due to the historical way

---

* Charles-Augustin de Coulomb (b. 1736 Angoulême, d. 1806 Paris), French engineer and physicist. His careful experiments on electric charges provided a firm basis for the study of electricity.

TABLE 3 Some observed electric fields.

OBSERVATION	ELECTRIC FIELD
Field 1 m away from an electron in vacuum	Challenge 9 s
Field values sensed by sharks	down to 0.1 μV/m
Cosmic noise	10 μV/m
Field of a 100 W FM radio transmitter at 100 km distance	0.5 mV/m
Field inside conductors, such as copper wire	0.1 V/m
Field just beneath a high power line	0.1 to 1 V/m
Field of a GSM antenna at 90 m	0.5 V/m
Field inside a typical home	1 to 10 V/m
Field of a 100 W bulb at 1 m distance	50 V/m
Ground field in Earth's atmosphere	100 to 300 V/m
Field inside thunder clouds	up to over 100 kV/m
Maximum electric field in air before sparks appear	1 to 3 MV/m
Electric fields in biological membranes	10 MV/m
Electric fields inside capacitors	up to 1 GV/m
Electric fields in petawatt laser pulses	10 TV/m
Electric fields in U^{91+} ions, at nucleus	1 EV/m
Maximum practical electric field in vacuum, limited by electron pair production	1.3 EV/m
Maximum possible electric field in nature (corrected Planck electric field $c^4/4Ge$)	$1.9 \cdot 10^{62}$ V/m

the unit of charge was defined first.* The essential point of the formula is the decrease of the field with the square of the distance; can you imagine the origin of this dependence? A simple way to picture Coulomb's formula is illustrated in Figure 10.

The two previous equations allow us to write the interaction between two charged bodies as

$$\frac{d\boldsymbol{p}_1}{dt} = \frac{1}{4\pi\varepsilon_0}\frac{q_1 q_2}{r^2}\frac{\boldsymbol{r}}{r} = -\frac{d\boldsymbol{p}_2}{dt} \ , \tag{5}$$

where $d\boldsymbol{p}$ is the momentum change, and $\boldsymbol{r}$ is the vector connecting the two centres of mass. This famous expression for electrostatic attraction and repulsion, also due to Coulomb, is valid only for charged bodies that are either of small size or spherical, and most of all, only for bodies that are *at rest* with respect to each other and to the observer. This description defines the field of *electrostatics*.

Electric fields accelerate charges. As a result, in everyday life, electric fields have two main properties: they contain energy and they can polarize bodies. The energy content

---

* Other definitions of this and other proportionality constants to be encountered later are possible, leading to *unit systems* different from the SI system used here. The SI system is presented in detail in Appendix A. Among the older competitors, the Gaussian unit system often used in theoretical calculations, the Heaviside–Lorentz unit system, the electrostatic unit system and the electromagnetic unit system are the most important ones.
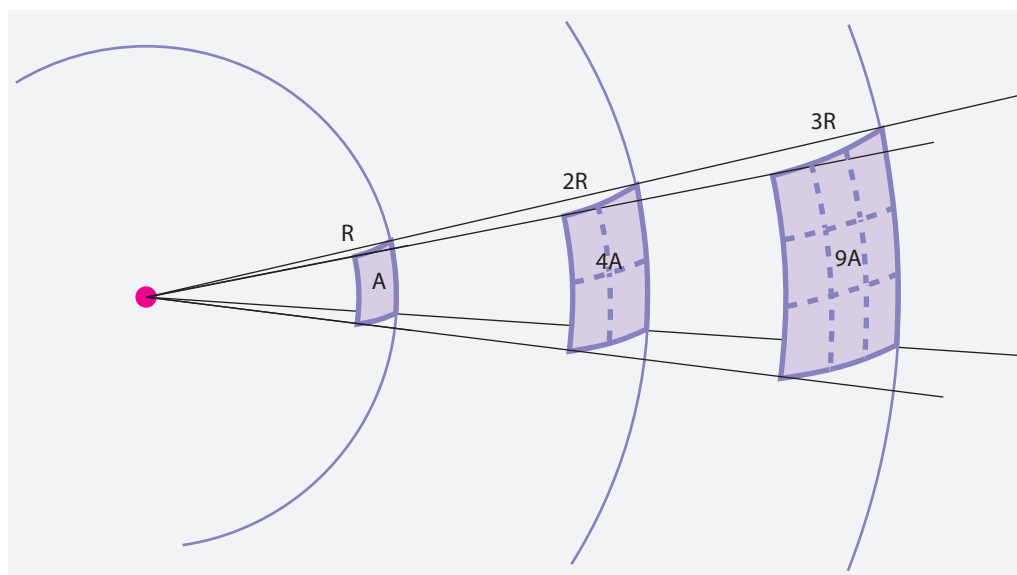
FIGURE 10 A visualization of Coulomb's formula and Gauss' law.

TABLE 4 Properties of the classical electric field: a (polar) vector at every point in space.

ELECTRIC FIELDS CAN	PHYSICAL PROPERTY	MATHEMATICAL NAME	DEFINITION
Attract bodies	accelerate charges	coupling	equation (4)
Repel bodies	accelerate charges	coupling	equation (4)
Be distinguished	distinguishability	element of set	Page 195
Change gradually	continuum	real vector space	Vol. I, page 73, Vol. V, page 284
Point somewhere	direction	vector space, dimensionality	Vol. I, page 73
Be compared	measurability	metricity	Vol. V, page 276
Be added	additivity	vector space	Vol. I, page 73
Have defined angles	direction	Euclidean vector space	Vol. I, page 73
Exceed any limit	infinity	unboundedness	Page 196
Change direction under reflection	polarity	parity-odd vector	
Keep direction under time reversal		time-even vector	

is due to the electrostatic interaction between charges. The strength of this interaction is considerable. For example, it is the basis for the force of our muscles. Muscular force is a macroscopic effect of Coulomb's relation (5). Another example is the material strength of steel or diamond. As we will discover, all atoms are held together by electrostatic at-

traction. To convince yourself of the strength of electrostatic attraction, answer the following: What is the force between two boxes with a gram of protons each, located on the two poles of the Earth? Try to guess the result before you calculate the astonishing value.

Coulomb's relation for the field around a charge can be rephrased in a way that helps to generalize it to non-spherical bodies. Take a closed surface, i.e., a surface than encloses a certain volume. Then the integral of the electric field over this surface, the electric flux, is the enclosed charge Q divided by ε_0:

$$\oint_{\text{closed surface}} \boldsymbol{E} \, \mathrm{d}\boldsymbol{A} = \frac{Q}{\varepsilon_0} \; . \tag{6}$$

This mathematical relation, called *Gauss's 'law'*,* from the result of Coulomb. (Note that in the simplified form given here, it is valid only for static situations.) Since inside conductors the electrical field is zero, Gauss's 'law' implies, for example, that if a charge q is surrounded by an uncharged metal sphere, the *outer* surface of the metal sphere shows the same charge q.

Do uncharged bodies attract one other? In first approximation they do not. But when the question is investigated more precisely, we will find that they can attract one other. Can you find the conditions for this to happen? In fact, the conditions are quite important, as our own bodies, which are made of neutral molecules, are held together in this way.

PUMPING CHARGE

Owing to the high strength of electromagnetic interactions, separating charges is not an easy task. This is the reason that electrical effects have only been commonly used for about a hundred years. Humanity had to wait for practical and efficient devices to be invented for separating charges and putting them into motion: to use electric effects, we need *charge pumps*. Some types are shown in Figure 11.

Of course, every charge pump requires energy. Batteries in mobile phones and the ion channels in living cells use chemical energy to do the trick. Thermoelectric elements, as used in some watches, use the temperature difference between the wrist and the air to separate charges; solar cells use light, and dynamos or Kelvin generators use kinetic energy. Can you explain whether batteries or any other of these devices sources of charges?

---

* Carl-Friedrich Gauß (b. 1777 Braunschweig, d. 1855 Göttingen), German mathematician. He was together with the Leonhard Euler, the most important mathematician of all times. A famous enfant prodige, when he was 19 years old, he constructed the regular heptadecagon with compass and ruler (see www.mathworld.wolfram.com/Heptadecagon.html). He was so proud of this result that he put a drawing of the figure on his tomb. Gauss produced many results in number theory, topology, statistics, algebra, complex numbers and differential geometry which are part of modern mathematics and bear his name. Among his many accomplishments, he produced a theory of curvature and developed non-Euclidean geometry. He also worked on electromagnetism and astronomy.

Gauss was a difficult character, worked always for himself, and did not found a school. He published little, as his motto was: pauca sed matura. As a consequence, when another mathematician published a new result, he regularly produced a notebook in which he had noted the very same result already years before. His notebooks are now available online at www.sub.uni-goettingen.de.

FIGURE 11 Various types of charge pumps: a bicycle dynamo, an alternator in a power station, a Windhurst machine, an electric eel, a voltaic cell, a leaf and a solar cell (© Wikimedia, Q-Cells).

WHAT IS ELECTRICITY?

The answer to this question is: *Electricity* is the name for a field of inquiry, but not the name for any specific observation or effect. Electricity is not a specific term; sometimes it is used to refer to electric current and its effects, sometimes to observations about of electric charge, sometimes to the effects of electric fields. In fact the vocabulary issue hides a deeper question that remains unanswered at the end of the twentieth century: what is the nature of electric charge? In order to solve this issue, we start with the following question.

CAN WE DETECT THE INERTIA OF ELECTRICITY?

If electric charge really is something *flowing* through metals, we should be able to observe the effects shown in Figure 12: electric charge should fall, should have inertia and should be separable from matter. Indeed, each of these effects has been observed. For

If electric charge in metals moves
like a fluid, it should:

fall under gravity

be subject to centrifugation

resist acceleration

lead to recoil just after
switching on a currrent

spray when pumped
strongly

prevent free charges
from falling through
a thin hollow tube
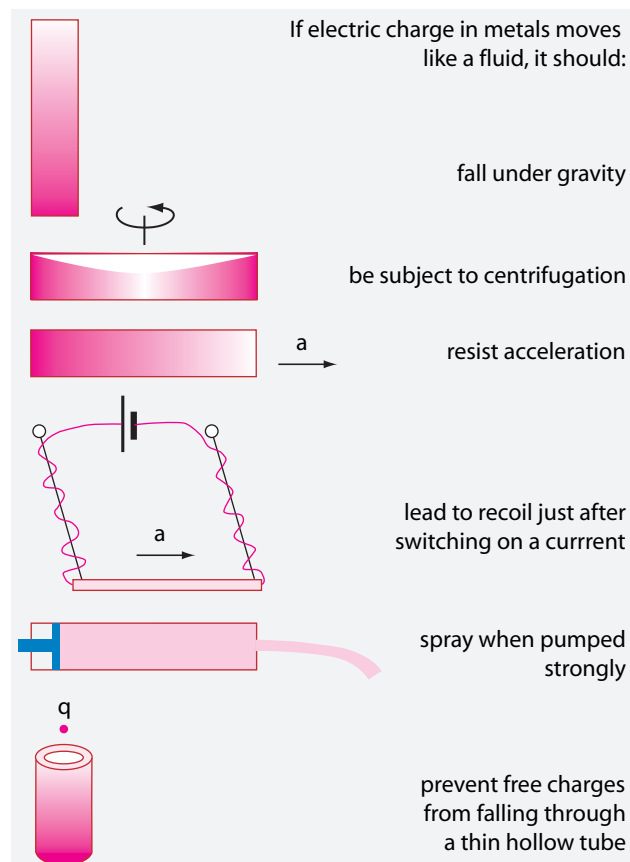
FIGURE 12 Consequences of the flow
of electricity.

example, when a long metal rod is kept vertically, we can measure an electrical potential
difference, a voltage, between the top and the bottom. In other words, we can measure
the *weight* of electricity in this way. Similarly, we can measure the potential difference
between the ends of an accelerated rod. Alternatively, we can measure the potential dif-
ference between the centre and the rim of a rotating metal disc. The last experiment was,
in fact, the way in which the ratio q/m for currents in metals was first measured with
precision. The result is

$$q/m \approx -1.8 \cdot 10^{11} \, \text{C/kg} \tag{7}$$

for all metals, with small variations in the second digit. The minus sign is due to the
definition of charge. In short, electrical charge in metals has mass, though a very small
one.

If electric charge has mass, whenever we switch on an electrical current, we get a *recoil*.
This simple effect can easily be measured and confirms the mass to charge ratio just given.
Also, the emission of current into air or into vacuum is observed; in fact, every television
tube uses this principle to generate the beam producing the picture. It works best for
metal objects with sharp, pointed tips. The rays created this way – we could say that they
are 'free' electricity – are called *cathode rays*. Within a few per cent, they show the same

Ref. 8

Ref. 9

Ref. 10

mass to charge ratio as expression (7). This correspondence thus shows that charges move almost as freely in metals as in air; this is the reason that metals are such good conductors.

If electric charge *falls* inside vertical metal rods, we can make the astonishing deduction that cathode rays – as we will see later, they consist of free electrons* – should not be able to fall through a vertical metal tube. This is due to exact compensation of the acceleration by the electrical field generated by the displaced electricity in the tube and the acceleration of gravity. Thus electrons should not be able to fall through a long thin cylinder. This would not be the case if electricity in metals did not behave like a fluid. The experiment has indeed been performed, and a reduction of the acceleration of free fall for electrons of 90 % has been observed. Can you imagine why the ideal value of 100 % is not achieved?

If electric current behaves like a liquid, we should be able to measure its speed. The first to do so, in 1834, was Charles Wheatstone. In a famous experiment, he used a wire of a quarter of a mile length, to produce three sparks: one at the start, one at the middle, and one at the end. He then mounted a rapidly moving mirror on a mechanical watch. By noting how much the three spark images were shifted against each other on a screen, he determined the speed to be 0.45 Gm/s, though with a large error. Latter, more precise measurements showed that the speed is always below 0.3 Gm/s, and that it depends on the metal and the type of insulation of the wire. The high value of the speed convinced many people to use electricity for transmitting messages. In fact, these experiments measure the signal speed of electromagnetic waves carried by metal wires. For the actual speed of electrc charges, see below. A modern version of the signal speed experiment, for computer fans, uses the 'ping' command from the UNIX operating system. The 'ping' command measures the time for a computer signal to reach another computer and return back. If the cable length between two computers is known, the signal speed can be deduced. Just try.

As a note, the speed of electricity is *too slow* for many people. Modern computers that are connected to stock exchanges are located as near as possible to the stock exchange, because the time advantage the short communication distance provides is essential for getting a good financial performance in certain trading markets.

Experiments with charges ejected from metals show that they have a charge to mass ratio of

$$q/m = -1.758\,820\,150(44) \cdot 10^{11}\ \text{C/kg} \tag{8}$$

The particles with this property are called *electrons*. Other types of charges, with different charge-to-mass ratio, also exist in nature. Examples are the *ions* found in batteries and leaves, the *muons* found in cosmic rays, and the *mesons* produced in particle accelerators. We will meet these particles later in our adventure.

In summary, experiments show that all charges have mass. And like all massive bodies, charges move slower than light. Charge is a property of matter; images and light have no

---

* The name 'electron' is due to George Stoney. Electrons are the smallest and lightest charges moving in metals; they are, usually – but not always – the 'atoms' of electricity – for example in metals. Their charge is small, 0.16 aC, so that flows of charge typical of everyday life consist of large numbers of electrons; as a result, electrical charge behaves like a continuous fluid. The particle itself was discovered and presented in 1897 by the Prussian physicist Johann Emil Wiechert (1861–1928) and, independently, three months later, by the British physicist Joseph John Thomson (1856–1940).

TABLE 5 Some observed electric current values.

OBSERVATION	CURRENT
Smallest current ever measured (for one moving electron)	3 aA
Human nerve signals	20 μA
Lethal current for humans	as low as 20 mA, typically 100 mA
Current drawn by a train engine	600 A
Current in a lightning bolt	10 to 100 kA
Highest current produced by humans	20 MA
Current inside the Earth, at the origin of its magnetic field	c. 100 MA
Maximum possible current in nature (corrected Planck electric current $e\sqrt{c^5/4\hbar G}$)	1.5 YA

TABLE 6 Some sensors for electrical current.

MEASUREMENT	SENSOR	RANGE
Conventional 20 euro multimeter	voltage drop over resistor	up to c. 3 A
Feeling threshold	human nerve	felt from 0.1 mA upwards
Reversible muscle contraction without danger	human nerve	up to 10 mA over long times, or up to 200 mA for at most 10 ms
Rhythm change	human heart	heart stops when about 20 mA flow through it
Strong muscle contraction with some damage	human nerve	up to 100 mA over long times, or up to 1 A for at most 200 ms
Smoke emission, strong burns	human flesh	from 1 A
Fire	trees	from 1 kA
Electric eel *Electrophorus electricus*	bult-in	up to 1 A and 500 V

charge.

FEELING ELECTRIC FIELDS

Why is electricity dangerous to humans? The main reason is that the human body is controlled by 'electric wires' itself. As a result, electricity applied to human bodies from the outside interferes with the internal signals. This has been known since 1789. In that year the Italian medical doctor Luigi Galvani (1737–1798) discovered that electrical current makes the muscles of a dead animal contract. The famous first experiment used frog legs: when electricity was applied to them, they twitched violently. Subsequent investigations confirmed that all nerves make use of electrical signals. Nerves are the 'control

TABLE 7 Searches for magnetic monopoles, i.e., for magnetic charges, in over 140 experiments.

SEARCH	MAGNETIC CHARGE
Smallest magnetic charge suggested by quantum theory	$g = \frac{h}{e} = \frac{eZ_0}{2\alpha} = 4.1\,\mathrm{pWb}$
Search in minerals, from mountains to the deep ocean	none, only dipoles Ref. 14
Search in meteorites and moon minerals	none, only dipoles Ref. 14
Search in cosmic rays	none, only dipoles Ref. 14
Search with particle accelerators	none, only dipoles Ref. 14

wires' of animals. We will explore nerves in more detail below.

Being electrically controlled, all mammals can sense strong electric fields. Humans can sense fields as low as $10\,\mathrm{kV/m}$, when hair stands on end. In contrast, several animals can sense much weaker electric (and magnetic) fields. Sharks, for example, can detect fields down to $1\,\mathrm{\mu V/m}$ using special sensors, the *Ampullae of Lorenzini*, which are found around their mouth. Sharks use them to detect the field created by prey moving in water; this allows them to catch their prey even in the dark. Several freshwater fish, the salamander and the platypus, the famous duck-billed mammal, can also sense electric fields. Like sharks, they use this ability to detect prey in water which is too muddy to see through. Certain fish, the so-called *weakly-electric fish*, even generate a weak field in order to achieve better prey detection.*

No land animal has special sensors for electric fields, because any electric field in air is strongly damped when it encounters a water-filled animal body. Indeed, the usual atmosphere has a low, vertical electric field of around $100\,\mathrm{V/m}$; inside the human body this field is damped to the $\mathrm{\mu V/m}$ range, which is much less than an animal's internal electric fields. In other words, humans do not have sensors for low electric fields because they are land animals. (Do humans have the ability to sense electric fields in water? Nobody seems to know.) However, there a few exceptions. You might know that some older people can sense approaching thunderstorms in their joints. This is due the coincidence between the electromagnetic field frequency emitted by thunderclouds – around $100\,\mathrm{kHz}$ – and the resonant frequency of nerve cell membranes.

The water content of the human body also means that the electric fields in air that are found in nature are rarely dangerous to humans. But whenever humans consciously sense electric fields, such as when high voltage makes their hair stand on end, the situation is potentially dangerous.

The high impedance of air also means that, in the case of time-varying electromagnetic fields, humans are much more prone to be affected by the magnetic component than by the electric component.

Magnets and other magnetic materials

The study of magnetism progressed across the world independently of the study of electricity. Towards the end of the twelfth century, the compass came into use in Europe. At

---

* It took until the year 2000 for technology to make use of the same effect. Nowadays, airbag sensors in cars often use electric fields to sense whether the person sitting in the seat is a child or an adult, thus changing the way that the airbag behaves in an accident.

FIGURE 13 Various types of magnets and effective magnets: the needle in a compass, some horseshoe magnets, two galaxies, the magnetic organ of a dove, the Earth, a lifting magnet, and the Sun. (© Wikimedia, Shambhavi, Anthony Ayiomamitis, NASA).

TABLE 8 Some observed magnetic fields.

OBSERVATION	MAGNETIC FIELD
Lowest measured magnetic field (e.g., fields of the Schumann resonances)	1 fT
Magnetic field produced by brain currents	0.1 pT to 3 pT
Magnetic field produced by single muscle action	1 pT
Intergalactic magnetic fields	1 pT to 10 pT
Magnetic field in the human chest, due to heart currents	100 pT
Magnetic field of our galaxy	0.5 nT
Magnetic field due to solar wind	0.2 to 80 nT
Magnetic field directly below high voltage power line	0.1 to 1 μT
Magnetic field of Earth	20 to 70 μT
Magnetic field inside home with electricity	0.1 to 100 μT
Magnetic field near mobile phone	100 μT
Magnetic field that influences visual image quality in the dark	100 μT
Magnetic field near iron magnet	100 mT
Solar spots	1 T
Magnetic fields near high technology permanent magnet	max 1.3 T
Magnetic fields that produces sense of coldness in humans	5 T or more
Magnetic fields in particle accelerator	10 T
Maximum static magnetic field produced with superconducting coils	22 T
Highest static magnetic fields produced in laboratory, using hybrid magnets	45 T
Highest *pulsed* magnetic fields produced without coil destruction	76 T
Pulsed magnetic fields produced, lasting about 1 μs, using imploding coils	1000 T
Field of white dwarf	10^4 T
Fields in petawatt laser pulses	30 kT
Field of neutron star	from 10^6 T to 10^{11} T
Quantum critical magnetic field	4.4 GT
Highest field ever measured, on magnetar and soft gamma repeater SGR-1806-20	0.8 to $1 \cdot 10^{11}$ T
Estimated magnetic field near atomic nucleus	1 TT
Maximum possible magnetic field in nature (corrected Planck magnetic field $c^3/4Ge$)	$6.3 \cdot 10^{53}$ T

that time, there were heated debates on whether it pointed to the north or the south. Then, in 1269, the French military engineer Pierre de Maricourt (1219–1292) published his study of magnetic materials. He found that every magnet has *two* points of highest magnetization, and he called them *poles*. He found that even after a magnet is cut, the resulting pieces always retain two poles: when the stone is left free to rotate, one points to the north and the other to the south. *All magnets are dipoles.* The two poles are called the
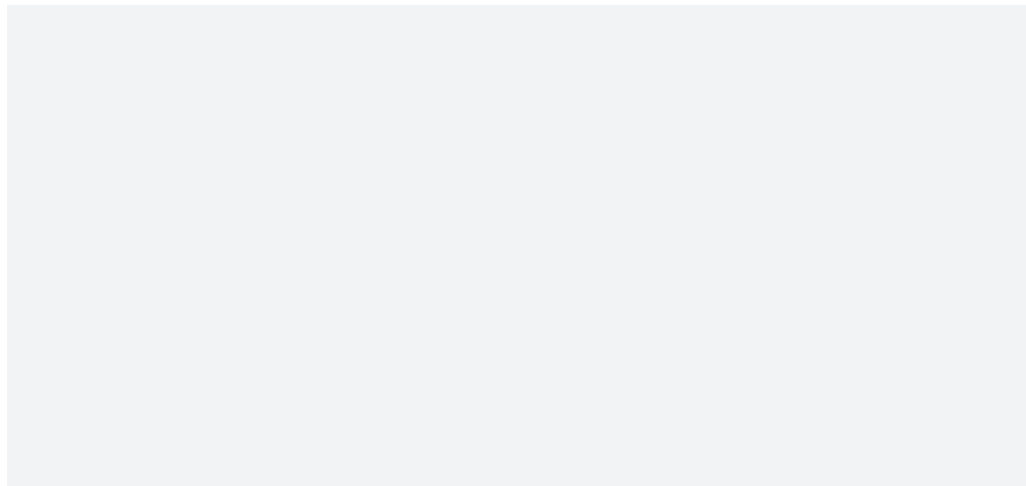
Ref. 15

FIGURE 14 Visualizing magnetic fields – with computer graphics and with iron filings.

magnet

diamagnetic
material

magnet

paramagnetic
material
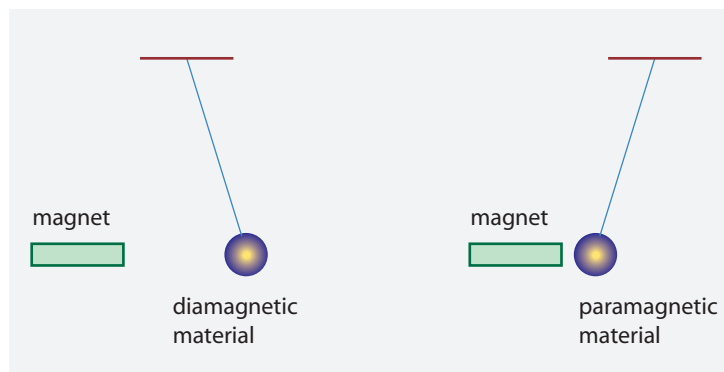
FIGURE 15 The two basic
types of magnetic material
behaviour (tested in an
inhomogeneous field):
diamagnetism and
paramagnetism.

north pole and the *south pole*. Like poles repel, and unlike poles attract. As a consequence, the magnetic north pole of the Earth is the one near the south pole, and vice versa.

Magnets are surrounded by magnetic fields; in other terms, they are surrounded by magnetic field lines. Magnetic fields, like electric fields, can be visualized with field lines. Figure 14 shows some ways to do this. We directly note the main difference between magentic and electric field lines: magnetic field lines have no beginning and no ends, they are closed. The direction of the field lines gives the direction of the magentic field, and the density of the lines gives the magnitude of the field.

Many systems in nature are magnets, as shown in Figure 13. The existence of two magnetic poles is valid for all magnets in nature: molecules, atoms and elementary particles are either dipoles or unmagnetic. There are no magnetic monopoles. Despite the promise of eternal fame, no magnetic monopole has ever been found, as summarized in Table 7.

Magnets have a second important property, shown in Figure 15: magnets transform unmagnetic materials into magnetic ones. There is thus a *magnetic polarization*, similar to the electric polarization. The amount of polarization depends on the material; some values are given in Table 9. Certain materials, the so-called *diamagnetic materials*, are

TABLE 9 The magnetic properties of materials.

MATERIAL	RELATIVE MAGNETIC PERMEABILITY μ_r
Diamagnetic materials $\mu_r < 1$, repelled by magnets	
Type I superconductors	0
Highly oriented pyrolitic graphite	0.999 55
Bismuth	0.999 83
Graphite	0.999 84
Gold	0.999 966
Copper	0.999 9936
Water	0.999 9912
Usual animals and plants	like water
Paramagnetic materials $\mu_r > 1$, attracted by magnets	
Air, oxygen	1.000 0019
Biomagnetic particles in living organisms	1.000 006
Aluminium	1.000 022
Platinum	1.000 26
Ferromagnetic materials $\mu_r \gg 1$, able to form magnets	
SmCo	*c.* 1.04
NdFeB	*c.* 1.15
Cobalt	80 to 200
Nickel	100
Iron	300 to 10 000
Permalloy	*c.* 8 000
Ferrites	up to 15 000
μ-metal	up to 140 000
Amorphous metals	up to 500 000

repelled by magnets, though usually by weak forces. Others, the so-called *paramagnetic* materials, are attracted to magnets. Some important materials, the *ferromagnetic materials*, such as steel, *retain* the induced magnetic polarization: they become permanently magnetized. This happens when the atoms in the material get aligned by an external magnet. Ferromagnetic materials are used to produce permanent *magnets* – thus artificial lodestone.

CAN HUMANS FEEL MAGNETIC FIELDS?

> “ Any fool can ask more questions than seven sages can answer. ”
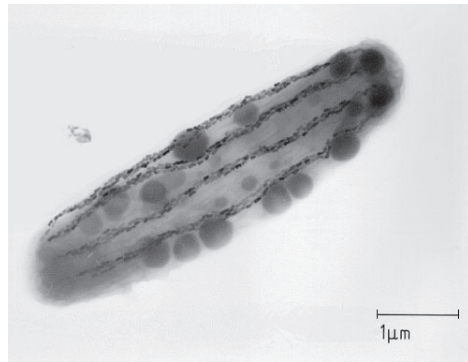>
> Antiquity

FIGURE 16 The magnetotactic bacterium *Magnetobacterium bavaricum* with its magnetosomes (© Marianne Hanzlik).

Ref. 16

It is known that honey bees, sharks, pigeons, salmon, trout, sea turtles, dolphins and certain bacteria can feel magnetic fields. One speaks of the ability for *magnetoreception*. All these life forms use this ability for navigation. The most common detection method is the use of small magnetic particles inside a cell; the cell then senses how these small built-in magnets move in a magnetic field. The magnets are tiny, typically around 50 nm in size. These small magnets are used to navigate along the magnetic field of the Earth. For higher animals, the variations of the magnetic field of the Earth, 20 to 70 μT, produce a landscape that is similar to the visible landscape for humans. They can remember it and use it for navigation.

Can humans feel magnetic fields? Magnetic material seems to be present in the human brain, but whether humans can feel magnetic fields is still an open issue. Maybe you can devise a way to check this?

Challenge 20 r

Magnetism and electricity

Are magnetism and electricity related? In the early 19th century, François Arago* discovered that they were. He explored a ship that had survived a bad thunderstorm. At that time, ships where made of wood. The ship had been struck by lightning; as a result, the ship needed a new compass. Thus lightning has the ability to demagnetize compasses. Arago knew that lightning is an electrical phenomenon. He concluded that magnetism and electricity must be related. More precisely, magnetism must be related to the *motion* of electricity.

If magnetism is related to motion of electricity, it must be possible to use magnetism and electricity to move matter.

How can one make a motor?

> " Communism is the power of the local councils
> plus electricification of the whole country.
>
> Lenin.**

---

* François Arago (1786–1853) French physicist.
** Lenin (b. 1870 Simbirsk, d. 1924 Gorki), founder of the Union of Soviet Socialist Republics, in 1920 stated this as the centre of his development plan for the country. In Russian, the local councils of that time were called soviets.
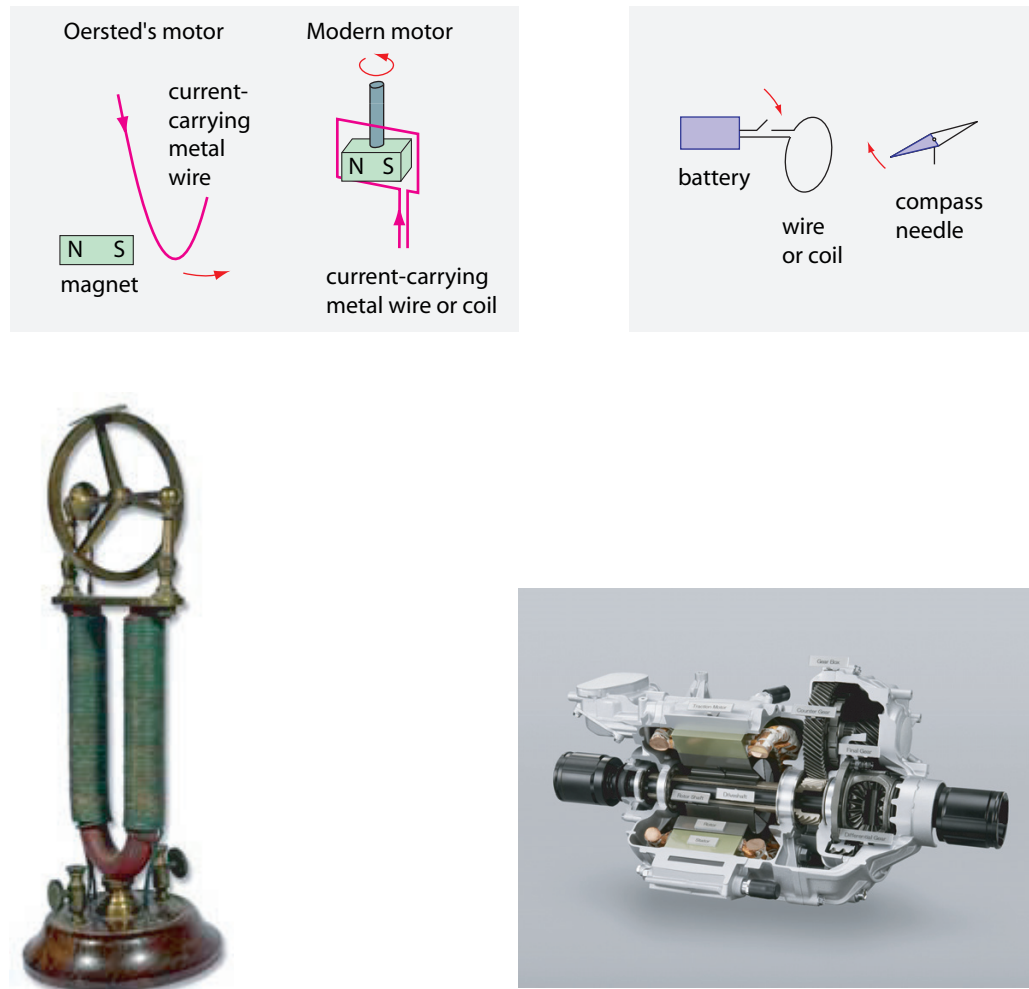
FIGURE 17 An old and a modern version of electric motor (© Wikimedia, Honda).

The reason for Lenin's famous statement were two discoveries. One was made in 1820 by Hans Christian Oersted* and the other in 1831 by Michael Faraday.** The consequences of these experiments changed the world completely in less than one century.

On the 21st of July of 1821, Hans Christian Oersted published a leaflet, in Latin, which took Europe by storm. Oersted had found (during a lecture demonstration to his stu-

---

* Hans Christian Oersted (1777–1851) Danish physicist.
** Michael Faraday (b. 1791 Newington Butts, d. 1867 London), English physicist, was born to a simple family, without schooling, and of deep and naive religious ideas. As a boy he became assistant to the most famous chemist of his time, Humphry Davy (1778–1829). He had no mathematical training, but late in his life he became member of the Royal Society. A modest man, he refused all other honours in his life. He worked on chemical topics, the atomic structure of matter and, most of all, he developed the idea of (magnetic) fields and field lines. He used fields to describe all his numerous experimental discoveries about electromagnetism, such as the Faraday effect. Fields were later described mathematically by Maxwell, who at that time was the only person in Britain to take over Faraday's field concept.

dents) that when a current is sent through a wire, a nearby magnet is put into motion. In other words, he found that the flow of electricity can *move* bodies.

Due to Oersted's leaflet, everybody in Europe with a bit of dexterity started to experiment with electricity. Further experiments show that *two* wires in which charges flow attract or repel each other, depending on whether the currents are parallel or antiparallel. These and other experiments show that wires in which electricity flows behave like magnets.* In other words, Oersted had found the definite proof that electricity could be turned into magnetism.

Shortly afterwards, Ampère** found that *coils* increase these effects dramatically. Coils behave like small magnets. In particular, coils, like magnets, always have two poles, usually called the north and the south pole. Opposite poles attract, like poles repel each other. As is well known, the Earth is itself a large magnet, with its magnetic north pole near the geographic south pole, and vice versa. However, the magnetic field of the Earth is *not* due to a solid permanent magnet inside it. The Earth's solid core is too hot to be a permanent magnet; instead, the magnetic field is due to circulating currents in the outer, liquid core.

All the relations between electricity and magnetism can be used to make electric motors. First, electric current is used to generate a magnetic field; then the field is used to move a magnet attached to the motor axis. The details on how to do this effectively depend on the size of the motor one is building, and form a science on its own: electric engineering. Figure 17 shows some examples of electric motors.

WHICH CURRENTS FLOW INSIDE MAGNETS?

Magnetic monopoles do not exist. Therefore, all magnetic fields in nature are due to moving electric charges. But that is strange: If all magnetic fields are due to the motion of charges, this must be also the case inside lodestone, or inside a usual permanent magnet. Can this be shown?

In 1915, two men in the Netherlands found a simple way to prove that in any permanent magnet, charges are moving. They suspended a metal rod from the ceiling by a thin thread and then put a coil around the rod, as shown in Figure 18. They predicted that the tiny currents inside the rod would become aligned by the magnetic field of the coil. As a result, they expected that a current passing through the coil would make the rod turn around its axis. Indeed, when they sent a strong current through the coil, the rod rotated. (As a result of the current, the rod was magnetized.) Today, this effect is called

Ref. 17

---

* In fact, if one imagines tiny currents moving in circles inside magnets, one gets a unique description for all magnetic fields observed in nature.

** André-Marie Ampère (b. 1775 Lyon, d. 1836 Marseille), French physicist and mathematician. Autodidact, he read the famous *Encyclopédie* as a child; in a life full of personal tragedies, he wandered from maths to chemistry and physics, worked as a school teacher, and published nothing of importance until 1820. Then the discovery of Oersted reached all over Europe: electrical current can deviate magnetic needles. Ampère worked for years on the problem, and in 1826 published the summary of his findings, which lead Maxwell to call him the 'Newton of electricity'. Ampère named and developed many areas of electrodynamics. In 1832, he and his technician also built the first *dynamo*, or rotative current generator. Of course, the unit of electrical current is named after him.

Ampère had two cats, which he liked dearly, a large one and a small one. When he was doing his experiments in his laboratory, they wanted to come in, and when they were in, they soon wanted to go out. One day he was fed up. He made two holes in his door, a large one and a small one.
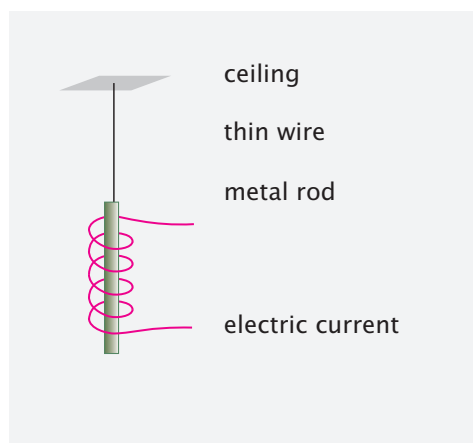
FIGURE 18 Current makes a metal rod rotate.

the *Einstein–de Haas effect* after the two men who imagined, measured and explained it.*
The effect thus shows that even in the case of a permanent magnet, the magnetic field is
due to the internal motion of charges. The magnitude of the effect also shows that the
moving particles are electrons. Twelve years later it became clear that the angular mo-
mentum responsible for the effect is a mixture of orbital and spin angular momentum;
in fact, the electron spin plays a central role in the effect.

Permanent magnets are made from ferromagnetic materials. Permanent magnetiza-
tion is due to the alignment of microscopic rotational motions. Due to this connection,
an even more surprising effect can be predicted: Rotating a piece of unmagnetized fer-
romagnetic material should magnetize it, because the tiny rotating currents would then
be aligned along the axis of rotation. This effect has indeed been observed; it is called
the *Barnett effect* after its discoverer. Like the Einstein–de Haas effect, the magnitude of
the Barnett effect can also be used to determine the gyromagnetic ratio of the electron.
Thus, also the Barnett effect proves that the spins of electrons (usually) play a larger role
in magnetism than their orbital angular momentum.

MAGNETIC FIELDS

All experiments show that the magnetic field has a given direction in space, and a magni-
tude common to all (resting) observers, whatever their orientation. We are thus tempted
to describe the magnetic field by a vector. However, this would be wrong, since a mag-
netic field does not behave like an arrow when placed before a mirror. Imagine that a
system produces a magnetic field directed to the right. You can take any system, a coil,
a machine, etc. Now build or imagine a second system that is the exact mirror version
of the first: a mirror coil, a mirror machine, etc. The magnetic system produced by the
mirror system does not point to the left, as maybe you expected: it still points to the right.
(Check by yourself.) In simple words, magnetic fields do *not* fully behave like arrows.

* Wander Johannes de Haas (1878–1960), Dutch physicist. De Haas is best known for two additional
magneto-electric effects named after him, the *Shubnikov–de Haas effect* (the strong increase of the magnetic
resistance of bismuth at low temperatures and high magnetic fields) and the *de Haas–van Alphen effect* (the
diamagnetic susceptibility of bismuth at low temperatures is a periodic function of the magnetic field).

TABLE 10 Properties of the classical magnetic field: an axial vector.

MAGNETIC FIELDS CAN	PHYSICAL PROPERTY	MATHEMATICAL NAME	DEFINITION
Attract currents	deflect charges	coupling	equation (10)
Repel currents	deflect charges	coupling	equation (10)
Be distinguished	distinguishability	element of set	Page 195
Change gradually	continuum	real vector space	Vol. I, page 73, Vol. V, page 284
Point somewhere	direction	vector space, dimensionality	Vol. I, page 73
Be compared	measurability	metricity	Vol. V, page 276
Be added	additivity	vector space	Vol. I, page 73
Have defined angles	direction	Euclidean vector space	Vol. I, page 73
Exceed any limit	infinity	unboundedness	Page 196
Keep direction under reflection	axiality	parity-even vector, pseudovector	
Change direction under time reversal	axiality	time-odd vector	

In other words, it is *not* completely correct to describe a magnetic field by a vector $B = (B_x, B_y, B_z)$, as vectors behave like arrows. The magnetic field is a *pseudovector*; angular momentum and torque are also examples of such quantities. The precise way is to describe the magnetic field by the quantity[*]

$$B = \begin{pmatrix} 0 & -B_z & B_y \\ B_z & 0 & -B_x \\ -B_y & B_x & 0 \end{pmatrix} , \tag{9}$$

called an *antisymmetric tensor*.

The *magnetic field* is defined by the acceleration it imparts on moving charges. This acceleration is observed to follow

$$a = \frac{e}{m} v \times B \tag{10}$$

a relation which is often called *Lorentz acceleration*, after the important Dutch physicist Hendrik A. Lorentz[**] who first stated it clearly.[***]

---

[*] The quantity B was not called the 'magnetic field' until recently. We follow here the modern, logical definition, which supersedes the traditional one, where B was called the 'magnetic flux density' or 'magnetic induction' and another quantity, H, was called – incorrectly, but for over a century – the magnetic field. This quantity H will not appear in this walk, but it is important for the description of magnetism in materials.

[**] Hendrik A. Lorentz, (b. 1853 Arnhem, d. 1928 Haarlem). For more details on his biography, see the volume on relativity.

[***] The expression $v \times B$ is the vector product of the two vectors. The most practical way to calculate the

TABLE 11 Some sensors for static and quasistatic magnetic fields.

MEASUREMENT	SENSOR	RANGE
Voltage	Hall probe	up to many T
Induced electromotive force (voltage)	doves	from a few nT
Bone growth stimulation	piezoelectricity and magnetostriction of bones	from 50 mT
Induced electromotive force (voltage)	human nerves	from a few T
Sensations in thorax and shoulders	human nerves	strong switched gradients
Sharks	induced voltage when waving left to right	a few nT
Plants	unclear	small effects on growth

The Lorentz acceleration is the effect at the root of any electric motor. An electric motor is a device that uses a magnetic field as efficiently as possible to accelerate charges flowing in a wire. Through the motion of the charges, the wire is then also moved. In an electric motor, electricity is thus transformed into magnetism and then into motion. The first efficient electric motors were built already in the 1830s.

Like for the electric field, we need to know how the *strength* of a magnetic field is determined by a moving charge. Experiments such as Oersted's show that the magnetic field is due to moving charges, and that a point-like charge moving with velocity $\boldsymbol{v}$ produces a field $\boldsymbol{B}$ given by

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} q \, \frac{\boldsymbol{v} \times \boldsymbol{r}}{r^3} \quad \text{where} \quad \frac{\mu_0}{4\pi} = 10^{-7} \, \text{N/A}^2 \,. \tag{13}$$

vector product $\boldsymbol{v} \times \boldsymbol{B}$ component by component is given by the determinant

$$\boldsymbol{v} \times \boldsymbol{B} = \begin{vmatrix} \boldsymbol{e}_x & v_x & B_x \\ \boldsymbol{e}_y & v_y & B_y \\ \boldsymbol{e}_z & v_z & B_z \end{vmatrix} \quad \text{or, more sloppily} \quad \boldsymbol{v} \times \boldsymbol{B} = \begin{vmatrix} + & - & + \\ v_x & v_y & v_z \\ B_x & B_y & B_z \end{vmatrix} \,. \tag{11}$$

This is easy to remember and easy to perform, both with letters and with numerical values. (Here, $\boldsymbol{e}_x$ is the unit basis vector in the x direction.) Written out, it is equivalent to the relation

$$\boldsymbol{v} \times \boldsymbol{B} = (v_y B_z - B_y v_z, B_x v_z - v_x B_z, v_x B_y - B_x v_y) \tag{12}$$

which is harder to remember.

The Lorentz relation is also called the *Laplace acceleration*. It defines the magnitude and the direction of the magnetic field $\boldsymbol{B}$. The unit of the magnetic field is called tesla and is abbreviated T. One has $1 \, \text{T} = 1 \, \text{N s/C m} = 1 \, \text{V s/m}^2 = 1 \, \text{V s}^2/\text{A m}$.

The definition of the magnetic field again assumes, like that of the electric field, that the test charge q is so small that it does not disturb the field B to be measured. Again, we ignore this issue, which means that we ignore all quantum effects, until later in our adventure.

The definition of the magnetic field also assumes that space-time is flat, and it ignores all issues due to space-time curvature.

Does the definition of magnetic field given here assume a charge speed much lower than that of light?

This is called *Ampère's 'law'*. Again, the strange factor $\mu_0/4\pi$ is due to the historical way in which the electrical units were defined. The constant μ_0 is called the *permeability of the vacuum* and is defined by the fraction of newtons per ampere squared given in the formula. It is easy to see that the magnetic field has an intensity given by vE/c^2, where E is the electric field measured by an observer moving *with* the charge. This is one of the many hints that magnetism is a relativistic effect.

We note that equation (13) is valid only for small velocities and accelerations. Can you find the general relation?

Electromagnetism

In 1831, Michael Faraday discovered an additional piece of the puzzle, one that even the great Ampère had overlooked. He found that a *moving* magnet could cause a current flow in an electrical circuit. Magnetism can thus be turned into electricity. This important discovery allowed the production of electrical current flow by generators, so-called *dynamos*, using water power, wind power or steam power. In fact, the first dynamo was built in 1832 by Ampère and his technician. Dynamos jump-started the use of electricity throughout the world. Behind every electrical wall plug there is a dynamo somewhere.

Oersted found that electric current can produce magnetic fields. Faraday found that magnetic fields could produce electric currents and electric fields. Electric and magnetic fields are two aspects of the same phenomenon: *electromagnetism*. It took another thirty years to unravel the full description.

Additional experiments show that magnetic fields also lead to electric fields when one changes to a moving viewpoint. You might check this on any of the examples of Figures 17 to 37. *Magnetism is relativistic electricity*. Electric and magnetic fields are partly transformed into each other when switching from one inertial reference frame to the other. Magnetic and electrical fields thus behave like space and time, which are also mixed up when changing from one inertial frame to the other. The theory of special relativity thus tells us that there must be a single concept, an *electromagnetic field*, describing them both. Investigating the details, one finds that the electromagnetic field F surrounding charged bodies has to be described by an antisymmetric 4-*tensor*

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix} \text{ or } F_{\mu\nu} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix} .$$

(14)

Obviously, the electromagnetic field F, and thus every component of these matrices, depends on space and time. The matrices show that electricity and magnetism are two faces of the same effect.* In addition, since electric fields appear only in the topmost row and leftmost column, the expressions show that in everyday life, for small speeds, electricity and magnetism *can* be separated. (Why?)

Using relativistic notation, the electromagnetic field is thus defined through the 4-

---

* Actually, the expression for the field contains everywhere the expression $1/\sqrt{\mu_0\varepsilon_0}$ instead of the speed of light c. We will explain the reason for this substitution shortly.

acceleration b that it produces on a charge q of mass m and 4-velocity u:

$$m\boldsymbol{b} = q\boldsymbol{F}\boldsymbol{u}$$

or, in 3-vector notation

$$\mathrm{d}E/\mathrm{d}t = q\boldsymbol{E}\boldsymbol{v} \quad \text{and} \quad \mathrm{d}\boldsymbol{p}/\mathrm{d}t = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \ . \quad (15)$$

The expressions show how the power $\mathrm{d}E/\mathrm{d}t$ and the three-force $\mathrm{d}\boldsymbol{p}/\mathrm{d}t$ depend on the electric and magnetic fields.* The 4-vector expression and the 3-vector expression describe the same content; the simplicity of the first one is the reason for the involved matrices (14) describing the electromagnetic field $\boldsymbol{F}$.

We stress that the extended *Lorentz relation* (15) is the *definition* of the electromagnetic field $\boldsymbol{F}$, since the field is defined as that 'stuff' which accelerates charges. In particular, all devices that put charges into motion, such as batteries and dynamos, as well as all devices that are put into motion by flowing charges, such as electric motors and muscles, are described by this relation. That is why this relation is usually studied, in the 3-vector form, already in secondary school. The Lorentz relation describes all cases in which the motion of objects can be seen by the naked eye or felt by our senses, such as the movement of an electrical motor in a high speed train, in a lift and in a dental drill, the motion of the picture generating electron beam in a television tube, or the travelling of an electrical signal in a cable and in the nerves of the body.

Ref. 19, Ref. 20

In summary, we found that the interaction between charges can be described in two statements: First, charges produce electric and magnetic fields; second, charges are affected by electric and magnetic fields.

Electromagnetic fields and Lagrangians

The electromagnetic field tensor $\boldsymbol{F}$ is an *antisymmetric* 4-tensor. (Can you write down the relation between $F^{\mu\nu}$, $F_{\mu\nu}$ and $F^{\mu}{}_{\nu}$?) Like any antisymmetric tensor, the electromagnetic field has two invariants, i.e., two deduced properties that are the same for every observer: the invariant expression

Challenge 26 ny

$$B^2 - E^2/c^2 = \frac{1}{2}\mathrm{tr}\,\boldsymbol{F}^2 \quad (17)$$

and the invariant product

$$4\boldsymbol{E}\boldsymbol{B} = -c\,\mathrm{tr}\,\boldsymbol{F}^*\boldsymbol{F} \ . \quad (18)$$

---

* In component notation, using the convention to sum over Greek indices that appear twice, the definition of the Lorentz force is

$$mb^{\mu} = m\frac{\mathrm{d}u^{\mu}}{\mathrm{d}\tau} = qF^{\mu}{}_{\nu}u^{\nu} \quad \text{or}$$

$$m\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{pmatrix} \gamma c \\ \gamma v_x \\ \gamma v_y \\ \gamma v_z \end{pmatrix} = q\begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{pmatrix}\begin{pmatrix} \gamma c \\ \gamma v_x \\ \gamma v_y \\ \gamma v_z \end{pmatrix} \ . \quad (16)$$

(Can you confirm this, using the definition of trace as the sum of the diagonal elements?)

The first invariant expression, $B^2 - E^2/c^2 = \frac{1}{2}\text{tr}\,\boldsymbol{F}^2$, turns out to be (proportional to) the Lagrangian density of the electromagnetic field. In particular, it is a scalar. This first invariant implies that if E is larger, smaller or equal to cB for one observer, it also is for all other observers. Like for all intensive quantities that evolve, the Lagrangian is proportional to the *square* of the intensive quantity. The minus sign in the expression is the same minus sign that appears also in $c^2 t^2 - x^2$: it results from the mixing of electric and magnetic fields by boosts.

The Lagrangian density can be used to define the classical action of the electromagnetic field:

$$S = \int \frac{\varepsilon_0}{2}E^2 - \frac{1}{2\mu_0}B^2 \, \mathrm{d}t\mathrm{d}V \;. \tag{19}$$

As usual, the action measures the change occurring in a system; it thus defines the amount of change that occurs when field lines move. As usual, the action can be used to describe the motion of the electromagnetic field by using the *principle of least action*. The principle of least action then implies the evolution equations of the electromagnetic field, which are called *Maxwell's field equations of electrodynamics*. This approach is the simplest way to deduce them. We will discuss the field equations in detail shortly.

The second invariant of the electromagnetic field tensor, $4\boldsymbol{EB} = -c\,\text{tr}\,\boldsymbol{F}^*\boldsymbol{F}$, is a pseudoscalar; it describes whether the angle between the electric and the magnetic field is acute or obtuse for all observers.[*]

THE USES OF ELECTROMAGNETIC EFFECTS

The application of electromagnetic effects to daily life has changed the world. For example, the installation of electric lighting in city streets has almost eliminated the previously so common night assaults. These and all other electrical devices exploit the fact that charges can flow in metals and, in particular, that electromagnetic energy can be transformed

— into mechanical energy – as done in loudspeakers, motors and muscles;
— into light – as in lamps, lasers and giant squids;
— into heat – as in electric ovens, blankets, tea pots and in the prey of electric eels;

---

[*] There is in fact a third Lorentz invariant, much less known. It is specific to the electromagnetic field and is a combination of the field and its vector potential:

$$\kappa_3 = \frac{1}{2}A_\mu A^\mu \boldsymbol{F}_{\rho\nu}\boldsymbol{F}^{\nu\rho} - 2A_\rho \boldsymbol{F}^{\rho\nu}\boldsymbol{F}_{\nu\mu}A^\mu$$

$$= (\boldsymbol{AE})^2 + (\boldsymbol{AB})^2 - |\boldsymbol{A} \times \boldsymbol{E}|^2 - |\boldsymbol{A} \times \boldsymbol{B}|^2 + 4\frac{\varphi}{c}(\boldsymbol{AE} \times \boldsymbol{B}) - \left(\frac{\varphi}{c}\right)^2 (E^2 + B^2) \;. \tag{20}$$

This expression is Lorentz (but not gauge) invariant; knowing it can help clarify unclear issues, such as the lack of existence of waves in which the electric and magnetic fields are parallel. Indeed, for plane monochromatic waves all three invariants *vanish* in the Lorentz gauge. Also the quantities $\partial_\mu J^\mu$, $J_\mu A^\mu$ and $\partial_\mu A^\mu$ are

Lorentz invariants. (Why?) The last one, the frame independence of the divergence of the four-potential, reflects the invariance of gauge choice. The gauge in which the expression is set to zero is called the *Lorentz gauge*.
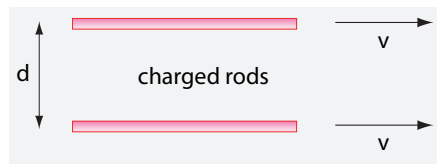
FIGURE 19 The relativistic aspect of magnetism.

— into chemical effects – as in hydrolysis, battery charging, electroplating and the brain;
— into coldness – as in refrigerators and Peltier elements, but in no known living system;
— into light signals – as in glass fibres, in glow worms and various deep ocean animals;
— into radio wave signals – as in radio and television, but in no known living system;
— into stored information – as in magnetic records, computers and human memory.

Electrical light, lasers, batteries, electric motors, radio, telephone, X-rays, television and computers have changed human life completely in less than one century.

We note that many of these uses also occur in biological systems. Could there be biological systems that communicate via radio waves?

How motors prove relativity to be right

> *The only mathematical operation I performed in my life was to turn the handle of a calculator.*
> Michael Faraday

All electric motors are based on the result that electric currents interact with magnetic fields. The simplest example is the attraction of two wires carrying parallel currents. This observation alone, made in 1820 by Ampère, is sufficient to make motion larger than a certain maximal speed impossible.

The argument is beautifully simple. We change the original experiment and imagine two long, electrically charged rods of mass m, moving in the same direction with velocity v and separation d. An observer moving with the rods would see an electrostatic repulsion between the rods given by

$$ma_e = -\frac{1}{4\pi\varepsilon_0}\frac{2\lambda^2}{d} \tag{21}$$

where λ is the charge per length of the rods. A second, *resting* observer sees two effects: the electrostatic repulsion and the attraction discovered by Ampère. The second observer therefore observes

$$ma_{em} = -\frac{1}{4\pi\varepsilon_0}\frac{2\lambda^2}{d} + \frac{\mu_0}{2\pi}\frac{\lambda^2 v^2}{d} \; . \tag{22}$$

This expression must be consistent with the observation of the first observer. This is the case only if both observers find repulsions. It is easy to check that the second observer

Challenge 29 ny

Ref. 22

Challenge 30 e

Challenge 31 e

sees a repulsion, as does the first one, only if

$$v^2 < \frac{1}{\varepsilon_0 \mu_0} = c^2 \; . \tag{23}$$

This maximum speed c, with a value of 0.3 GM/s, is thus valid for any object carrying charges. But *all* everyday objects contain charges: there is thus a maximum speed for matter.

Are you able to extend the argument for a maximum speed to neutral particles as well? We will find out more on this limit velocity, which we know already, in a minute.

Another argument for magnetism as a relativistic effect is the following. In a wire with electrical current, the charge is zero for an observer at rest with respect to the wire. The reason is that the charges enter and exit the wire at the same time for that observer. Now imagine an observer who flies along the wire. The entrance and exit events do not occur simultaneously any more; the wire is charged for a moving observer. (The charge depends on the direction of the observer's motion.) In other words, if the observer himself were charged, he would experience a force. Moving charges experience forces from current-carrying wires. This is exactly why magnetic fields were introduced: they only produce forces on *moving* charges. In short, current carrying wires are surrounded by magnetic fields.

In summary, electric effects are due to flow of electric charges and to electric fields; magnetism is due to *moving* electric charges.* Magnetism is *not* due to magnetic charges: as mentioned, such magnetic monopoles do not exist. The strength of magnetism, used in any running electric motor, proves relativity right: there is a maximum speed in nature for all masses and charges. Both electric and magnetic fields carry energy and momentum. They are two faces of the same coin.

CURIOSITIES AND FUN CHALLENGES ABOUT THINGS ELECTRIC AND MAGNETIC

> ❝ Et facta mirari et intellectua assequi. ❞
>
> Augustine of Hippo

Before we study the motion of an electromagnetic field in detail, let's have some fun with electricity.

* *

Nowadays, having fun with sparks is straightforward. Tesla coils, named after Nikola Tesla** are the simplest devices that allow long sparks to be produced at home. Atten-

---

* 'Electrons move in metal with a speed of about 1 μm/s; thus if I walk with the same speed along a cable carrying a constant current, I should not be able to sense any magnetic field.' What is wrong with this argument?

** Никола Тесла (1856 Smiljan–1943 New York City), Serbian engineer and inventor. He invented and promoted the polyphase alternating current system, the alternating current electric motor, wireless communication, fluorescent lighting and many other applications of electricity. He is also one of the inventors of radio. The SI unit of the magnetic field is named after him. A flamboyant character, his ideas were sometimes unrealistic; for example he imagined that Tesla coils could be used for wireless power transmission.

TABLE 12 Voltage values observed in nature.

OBSERVATION	VOLTAGE
Smallest measured voltage	*c.* 10 fV
Human nerves	70 mV
Volta cell	1 V
Voltaic cell ('battery')	1.5 V
Mains in households	230 V or 110 V
Electric eel	100 to 600 V
Tramway supply	500 V
Sparks when rubbing a polymer pullover	1 kV
Electric fence	0.7 to 10 kV
Train supply	15 kV
Ignition plug in cars	15 kV
Colour television tube	30 kV
X-ray tube	30 to 200 kV
Electron microscopes	0.5 kV to 3 MV
Stun gun	65 to 600 kV
Lightning stroke	10 to 100 MV
Record accelerator voltage	1 TV
Maximum possible voltage in nature (corrected Planck voltage $\sqrt{\hbar c^5/4G}/e$)	$6.1 \cdot 10^{27}$ V

tion: this is dangerous; that is the reason that such devices cannot be bought (almost) anywhere. The basic diagram and an example is shown in Figure 20. Tesla coils look like large metal mushrooms (to avoid unwanted discharges) and plans for their construction can be found on numerous websites or from numerous enthusiast's clubs, such as www. stefan-kluge.de.

* *

In 1722, George Graham discovered, by watching a compass needle, that the magnetic field of the Earth shows daily variations. Can you imagine why these variations occur?

Challenge 34 s

* *

If even knocking on a wooden door is an electric effect, we should be able to detect fields when doing so. Can you devise an experiment to check this?

Challenge 35 d

* *

Birds come to no harm when they sit on unprotected electricity lines. Nevertheless, one almost never observes any birds on tall, high voltage lines of 100 kV or more, which transport power across longer distances. Why?

Challenge 36 s

* *

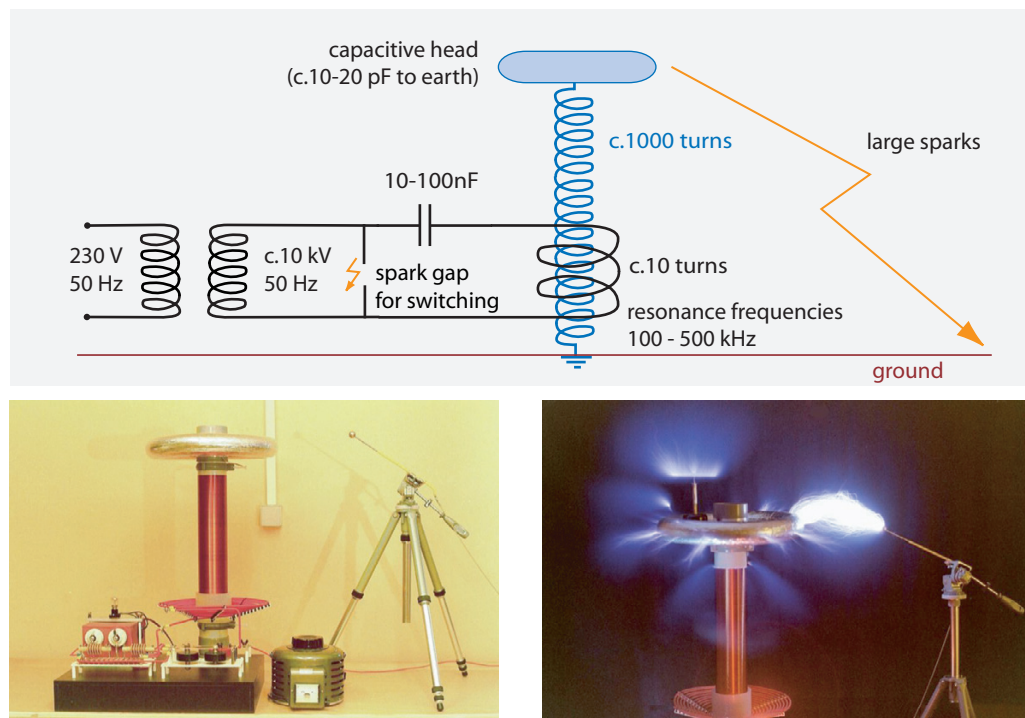How can you distinguish a magnet from an unmagnetized metal bar of the same size and

FIGURE 20 The schematics, the realization and the operation of a Tesla coil, including spark and corona discharges (photographs © Robert Billon).

Challenge 37 s   material, using no external means?

* *

In the basement of a house there are three switches that control three light bulbs in the first floor. You are in the basement and are allowed to go to the first floor only once. How Challenge 38 s do you find out which switch controls which bulb?

* *

How do you wire up a light bulb to the mains and three switches so that the light can be switched on at any of the switches and off at any other switch? And for four switches? Challenge 39 s Nobody will take a physicist seriously who is able to write Maxwell's equations but cannot solve this little problem.

* *

The first appliances built to generate electric currents were large rubbing machines. Then, in 1799 the Italian scientist Alessandro Volta (1745–1827) invented a new device to generate electricity and called it a *pile*; today its basic element is called a *(voltaic) cell*, a *primary cell*[*] or, less correctly, a *battery*. (Correctly speaking, a battery is a collection of cells, as the one found in a car.) Voltaic cells are based on chemical processes; they provide much

---

[*] A *secondary cell* is a rechargeable cell.

FIGURE 21 A common playground effect (© Evan Keller).

more current and are smaller and easier to handle than electrostatic machines. The invention of the battery changed the investigation of electricity so profoundly that Volta became world famous. At last, a simple and reliable source of electricity was available for use in experiments; unlike rubbing machines, piles are compact, work in all weather conditions and make no noise.

An apple or a potato or a lemon with a piece of copper and one of zinc inserted is one of the simplest possible voltaic cells. It provides a voltage of about 1 V and can be used to run digital clocks or to produce clicks in headphones. Volta was also the discoverer of the charge law $q = CU$ of capacitors (C being the capacity, and U the voltage) and the inventor of the high sensitivity capacitor electroscope. A modest man, nevertheless, the unit of electrical potential, or 'tension', as Volta used to call it, was deduced from his name. A 'battery' is a large number of voltaic cells; the term was taken from an earlier, almost purely military use.* A battery in a mobile phone is just an elaborated replacement for a number of apples or potatoes.

* *

What happened in Figure 21? Why are most of such pictures taken in good weather and with blond children?

* *

A PC or a telephone can communicate without wires, by using radio waves. Why are these and other electrical appliances not able to obtain their *power* via radio waves, thus eliminating power cables?

* *

Magnets can be used to accelerate steel balls. The most famous example is the *Gauss rifle*

---

* A pile made of sets of a zinc plate, a sheet of blotting paper soaked with salt water and a copper coin is easily constructed at home and tested with a calculator or a digital watch.
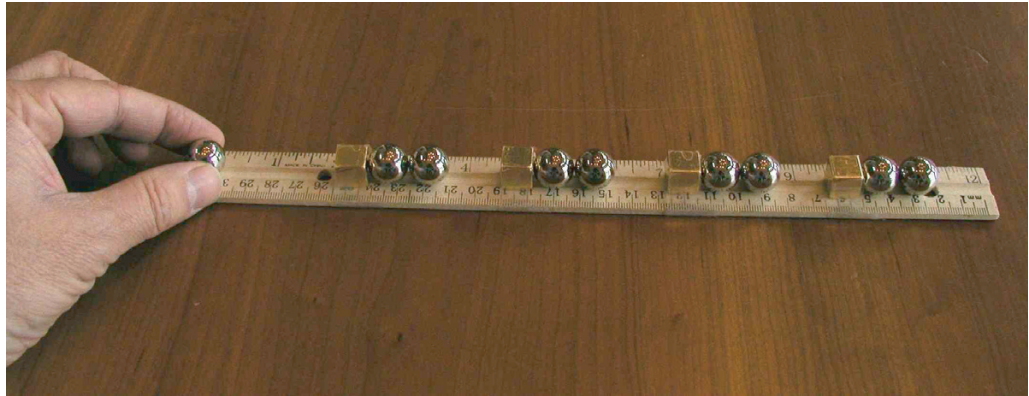
FIGURE 22 A Gauss rifle, made with a few steel balls and four magnets attached to a ruler with scotch tape (© Simon Quellen Field).

shown in Figure 22. If the leftmost ball is gently rolled towards the first magnet, the third ball is strongly kicked away. Then the process repeats: the speed increases even more for the fifth, the seventh and the ninth ball. The experiment never fails to surprise whoever sees it for the first time. Where does the momentum of the final ball come from?

Challenge 43 e

* *

Objects that are not right–left symmetric are called *chiral*, from the Greek word for 'hand'. Can you make a mirror that does not switch chirality (i.e., does not 'switch left and right')? In two different ways?

Challenge 44 s

* *

A Scotch tape roll is a dangerous device. Pulling the roll quickly leads to light emission (through triboluminescence) and to small sparks. It is suspected that several explosions in mines were triggered when such a spark ignited a combustible gas mixture.

* *

Take an envelope, wet it and seal it. After letting it dry for a day or more, open it in the dark. At the place where the two sides of paper are being separated from each other, the envelope glows with a blue colour. Why? Is it possible to speed up the test using a hair dryer?

Challenge 45 s

* *

A charge in an electric field feels a force. In other words, electric field produce a *potential energy* for charges. Since energy is conserved, electric potential energy can be transformed into kinetic energy or in thermal energy. What do these possibilities allow doing? What do they prevent from doing?

Challenge 46 e

* *

Electromagnetism is full of surprises and offers many effects that can be reproduced at home. The internet is full of descriptions of how to construct Tesla coils to produce
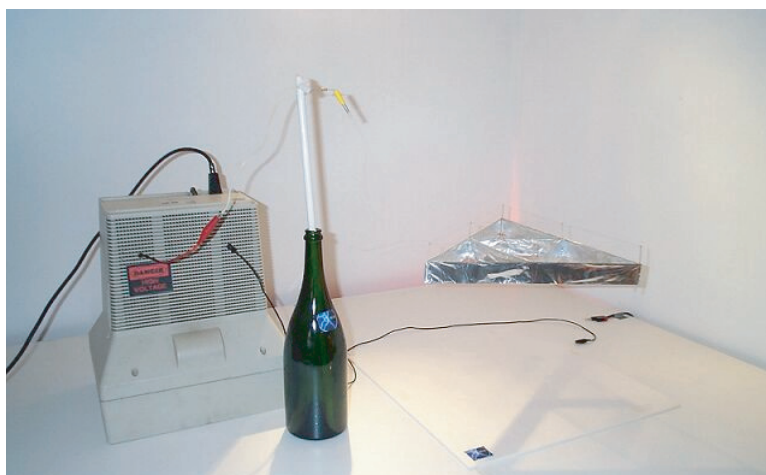
FIGURE 23 Lifting a light object – covered with aluminium foil – using high a tension discharge (© Jean-Louis Naudin at www.jlnlabs.org).

sparks, coil guns or rail guns to shoot objects, electrostatic machines to make your hair stand on end, glass spheres with touch-sensitive discharges and much more. If you like experiments, just search for these terms.

* *

A high voltage can lead to current flow through air, because air becomes conductive in high electric fields. In such discharges, air molecules are put in motion. As a result, one can make objects that are attached to a pulsed high tension source lift up in the air, if one optimizes this air motion so that it points downwards everywhere. The high tension is thus effectively used to accelerate ionized air in one direction and, as a result, an object will move in the opposite direction, using the same principle as a rocket. An example is shown in Figure 23, using the power supply of a PC monitor. (Watch out: danger!) Numerous websites explain how to build these so-called lifters at home; in Figure 23, the bottle and the candle are used as high voltage insulator to keep one of the two thin high voltage wires (not visible in the photograph) high enough in the air, in order to avoid discharges to the environment or to interfere with the lifter's motion. Unfortunately, the majority of websites – not all – give incorrect or confused explanations of the phenomenon. These websites thus provide a good challenge for one to learn to distinguish fact from speculation.

Challenge 47 e

* *

The electric effects produced by friction and by liquid flow are usually small. However, in the 1990s, a number oil tankers disappeared suddenly. The sailors had washed out the oil tanks by hosing sea water onto the tank walls. The spraying led to charging of the tank; a discharge then led to the oil fumes in the tank igniting. This led to an explosion and subsequently the tankers sank. Similar accidents also happen regularly when chemicals are moved from one tank to another.
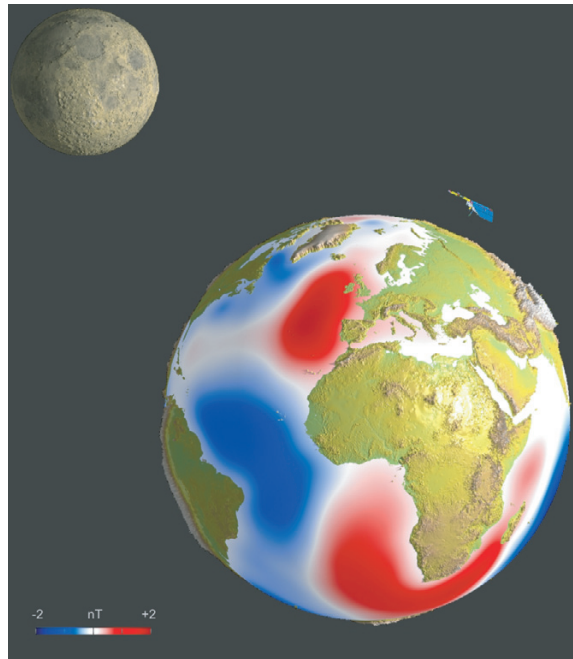
* *

FIGURE 24 The magnetic field due to the tides (© Stefan Maus).

Rubbing a plastic spoon with a piece of wool charges it. Such a charged spoon can be used to extract pepper from a salt–pepper mixture by holding the spoon over the mixture. Why?

Challenge 48 s

* *

When charges move, they produce a magnetic field. In particular, when ions inside the Earth move due to heat convection, they produce the Earth's magnetic field. When the ions high up in the atmosphere are moved by solar wind, a geomagnetic storm appears; its field strength can be as high as that of the Earth itself. In 2003, an additional mechanism was discovered. When the tides move the water of the oceans, the ions in the salt water produce a tiny magnetic field; it can be measured by highly sensitive magnetometers in satellites orbiting the Earth. After two years of measurements from a small satellite it was possible to make a beautiful film of the oceanic flows. Figure 24 gives an impression.

Ref. 23

* *

The names electrode, electrolyte, ion, anode and cathode were suggested by William Whewell (1794–1866) on demand of Michael Faraday; Faraday had no formal education and asked his friend Whewell to form two Greek words for him. For anode and cathode, Whewell took words that literally mean 'upward street' and 'downward street'. Faraday then popularized these terms, like the other words mentioned above.

* *

The shortest light pulse produced so far had a duration of 100 as. To how many wave-
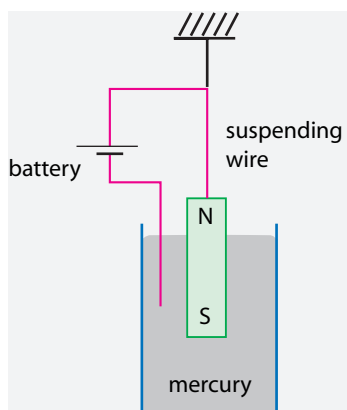
FIGURE 25 A unipolar motor.
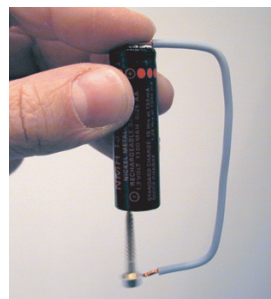
FIGURE 26 The simplest motor (© Stefan Kluge).

Challenge 49 s lengths of green light would that correspond?

* *

How long can batteries last? At Oxford University, in Clarendon Hall, visitors can watch a battery-operated electric bell that is ringing since 1840. The two batteries, two Zamboni piles, produce a high voltage and low current, sufficient to keep the bell ringing. Several other similar devices, using Zamboni piles, have worked in Italy with the same batteries for over 100 years.

* *

Challenge 50 s Why do we often see shadows of houses and shadows of trees, but never shadows of the electrical cables hanging over streets?

* *

Challenge 51 s How would you measure the speed of the tip of a lightning bolt? What range of values do you expect?

* *

Ref. 24 One of the simplest possible electric motors was discovered by Faraday in 1831. A magnet suspended in mercury will start to turn around its axis if a current flows through it. (See Figure 25.) In addition, when the magnet is forced to turn, the device (often also called Barlow's wheel) also works as a current generator; people have even tried to generate Challenge 52 s domestic current with such a system! Can you explain how it works?

The modern version of this motor makes use of a battery, a wire, a conductive samarium–cobalt magnet and a screw. The result is shown in Figure 26.

* *

Ref. 25 The magnetic field of the Earth has a dipole strength of $7.8 \cdot 10^{22}$ A m^2. It shields us, to-gether with the atmosphere, from lethal solar winds and cosmic radiation particles, by deflecting them to the poles. Today, a lack of magnetic field would lead to high radiation
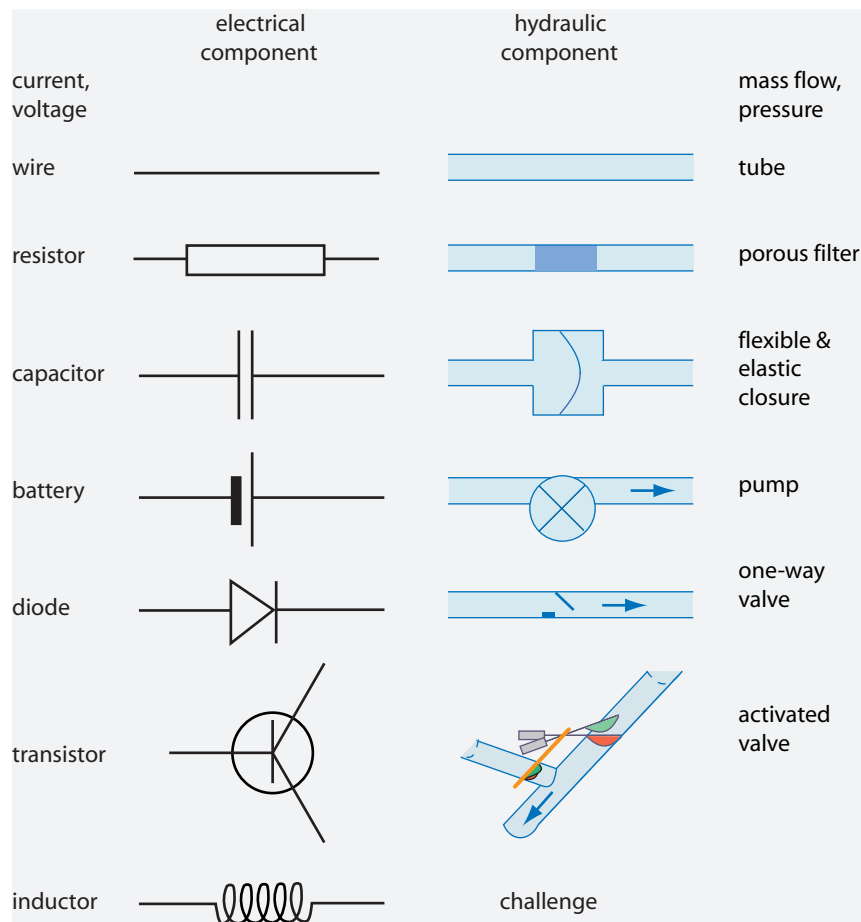
	electrical component	hydraulic component	
current, voltage			mass flow, pressure
wire			tube
resistor			porous filter
capacitor			flexible & elastic closure
battery			pump
diode			one-way valve
transistor			activated valve
inductor		challenge	

FIGURE 27 The correspondence of electronics and water flow.

on sunny days; but in the past, its lack would have prevented the evolution of the human species. We owe our existence to the magnetic field.

∗ ∗

Comparing electricity with water is a good way of understanding electronics. Figure 27 shows a few examples that even a teenager can use. Can you fill in the correspondence for the coil, and thus for a transformer?

Challenge 53 s

The picture also includes the *transistor*. This device, as the hydraulic component shows, can be used to control a large current by using a small current. Therefore, transistors can be used as *switches* and as *amplifiers*. This is the reason that all electronic circuits, from radios to mobile phones and computers – make heavy use of transistors. A modern mobile phone or computer typically contains several million transistors, mostly assembled inside so-called *integrated circuits*. The design of these devices is a science on its own.

∗ ∗

There is even a way to push the previous analogy in another direction: it is possible to produce a mathematically consistent analogy between electric circuits and continuous fields. The required circuits are infinite grids or meshes in all directions of space, and are called *mimetic discretizations*. If you like to think in electric terms, you might enjoy pursuing this. Just search for the term in a search engine.

* *

The ionosphere around the Earth has a resonant frequency of 7 Hz; for this reason any apparatus measuring low frequencies always gets a strong signal at this value. Can you give an explanation of the frequency?

* *

The Kirlian effect, which allows one to make such intriguingly beautiful photographs, is not a property of objects, but a result of the applied time-varying electric field.

* *

At home, electricity is mostly used as alternating current. In other words, no electrons actually flow through cables; as the drift speed of electrons in copper wires is of the order of 1 μm/s, electrons just move back and forward by 20 nm. Nothing flows in or out of the cables! Why do the electricity companies require a real flow of money in return, instead of being satisfied with a back and forth motion of money?

* *

Do electrons and protons have the same charge? Experiments show that the values are equal to within at least twenty digits. How would you check this?

* *

Charge is also velocity-independent. How would you check this?

* *

Magnets can be used, even by school children, to climb steel walls. Have a look at the www.physicslessons.com/TPNN.htm website.

* *

Can magnets be used to make a floating bed? In 2006, a Dutch architect presented to the public a small model of a beautiful floating bed, shown on the left of Figure 28, kept floating in the air by permanent magnets. To prevent that the model bed falls over, it is fastened to the ground by four ropes. On his website, the architect also offers a real-size version of the same bed, for a price of over one million US dollars. However, the images of the scaled up bed – the only two images that exist – are not photographs, but computer graphics, as this dream bed is impossible. Why?

* *

Extremely high magnetic fields have strange effects. At fields of 10^{10} T, vacuum becomes effectively birefringent, photons can split and coalesce, and atoms get squeezed. Hydrogen atoms, for example, are estimated to get two hundred times narrower in one direc-

FIGURE 28 The floating bed problem: while the left model, with a length of around 40 cm and a floating height of a few centimetres, exists and has been admired by many, the scaled-up, real-size version on the right is impossible (© Janjaap Ruissenaars at www.UniverseArchitecture.com). The two images on the right are *not* photographs: they show a dream, not reality. Why?

tion. Fortunately, these conditions exist only in specific neutron stars, called *magnetars*.

∗ ∗

Ohm's 'law', the observation that for almost all materials the current I is proportional to the voltage U, is

$$U \sim I \quad \text{or} \quad \frac{U}{I} = R = \text{const.} \tag{24}$$

and is due to a school teacher. Georg Simon Ohm (b. 1789 Erlangen, d. 1854 München), was a Bavarian school teacher and physicist. He explored the validity of the proportionality in great depth and for many materials; in those days, such measurements were difficult to perform. Ohm discovered that the proportionality applies to most materials and to many current levels, as long as the temperature, the material density and the charge densities remain constant. The proportionality is thus not valid for situations with sparks or in semiconductors. But it is valid for most solid conductors, in particular for metals.

Ohm's efforts were recognized only late in his life, and he eventually was promoted to professor at the Technical University in Munich. Later the unit of *electrical resistance R* – this is the official name for the proportionality factor between voltage, which is measured in *volt*, and current, which measured in *ampere* – was named after him. One *ohm* is defined and written as 1 V/A=1 Ω.

Today, Ohm's relation is easy to measure. Recently, even the electrical resistance of single atoms has been measured: in the case of xenon it turned out to be about 10^5 Ω. It was also found that lead atoms are ten times more conductive than gold atoms. Can you imagine why?

$$e^{\frac{\pi d U_{34}}{I_{12}\rho}} + e^{\frac{\pi d U_{41}}{I_{23}\rho}} = 1 \qquad\qquad \rho = \pi d U_{34}/(I_{12}\ln 2)$$

material thickness d
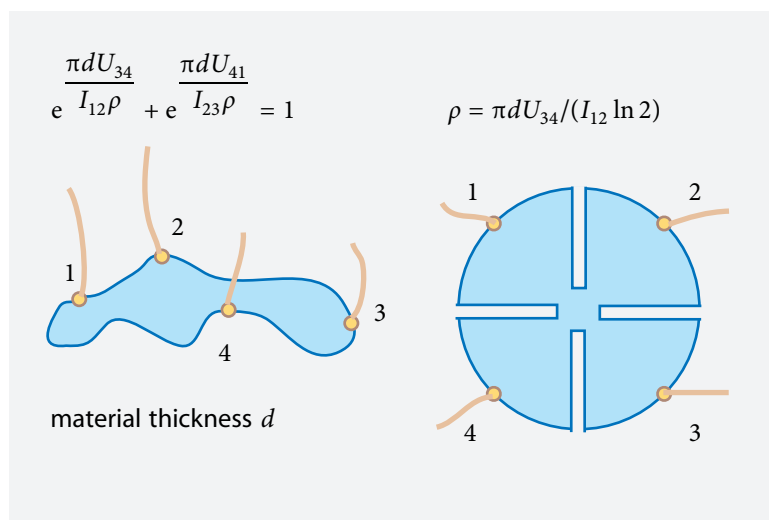
FIGURE 29 Can you deduce Van der Pauw's formula for the specific resistance ρ of homogeneous layers of any shape (left) or its special case for a symmetrical shape (right)?

* *

Since many decades, Ohm's 'law' is taught in secondary school until every pupil in a class has lost his interest in the matter. For example, the electric power P transformed into heat in a resistor is given

$$P = UI = I^2 R = \frac{U^2}{R} \; . \tag{25}$$

We mentioned this relation already earlier on; have a look. Now you know everything that needs to be known on the topic. Above all, the expression for electric power in a resistor describes electric heating, for example the heating in a modern kitchen stove or in a coffee machine.

* *

Ohm's 'law', so simple it seems, has many fascinating mathematical aspects. For example, in 1958, the Dutch physicist J.L. van der Pauw proved an astonishing formula and method that allows measuring the specific resistance ρ of material layers of *any* shape. One only needs to attach four gold wires to the layer anywhere on its border. The specific resistance is then given by the expression shown in Figure 29. Can you imagine how the formula is deduced? (This is not an easy problem.) The formula reduced the workload in laboratories across the world by a significant amount; before the formula had been discovered, in every experiment, researchers also had to produce separate samples that allowed to measure the specific resistance of the material they were investigating.

* *

A good way to make money is to produce electricity and sell it. In 1964, a completely new method was invented by Fletcher Osterle. The method was presented to a larger public in a beautiful experiment in 2003. One can take a plate of glass, add a conducting layers on each side, and then etch a few hundred thousand tiny channels through the plate, each
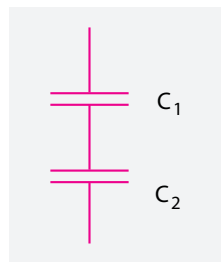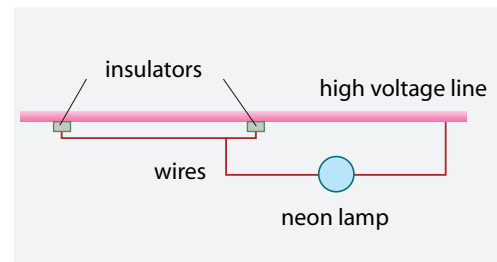
FIGURE 30
Capacitors in series.

FIGURE 31 A neon lamp hanging from a high
voltage line.

around 15 μm in diameter. When water is made to flow through the channels, a current
is generated. The contacts at the two conducting plates can be used like battery contacts.
    This simple device uses the effect that glass, like most insulators, is covered with a
charged layer when it is immersed in a liquid. Can you imagine why a current is gener-
Challenge 61 s   ated? Unfortunately, the efficiency of electricity generation is only about 1%, making the
method much less interesting than a simple blade wheel powering a dynamo.

* *

For beautiful animations about magnetic and electric fields, see the website web.mit.edu/
8.02t/www/802TEAL3D/visualizations.

* *

Gallium arsenide semiconductors can be patterned with so-called quantum dots and
point contacts. These structures allow one to count single electrons. This is now routinely
done in several laboratories around the world.

* *

Ref. 28   The charges on two capacitors in series are not generally equal, as naive theory states.
For perfect, leak-free capacitors the voltage ratio is given by the inverse capacity ratio
$V_1/V_2 = C_2/C_1$, due to the equality of the electric charges stored. This is easily deduced
from Figure 30. However, in practice this is only correct for times between a few and a
Challenge 62 s   few dozen minutes. Why?

On certain high voltage cables leading across the landscape, small neon lamps shine
when the current flows, as shown in Figure 31. (You can see them from the train when
Challenge 63 ny   riding from Paris to the Roissy airport.) How is this possible?

* *

Electric polarizability is the property of matter responsible for the deviation of water flow-
Page 16   ing from a tap caused by a charged comb. It is defined as the strength of electric dipole
induced by an applied electric field. The definition simply translates the observation that
many objects acquire a charge when an electric field is applied. Incidentally, how pre-
cisely combs get charged when rubbed, a phenomenon called *electrification*, is still one
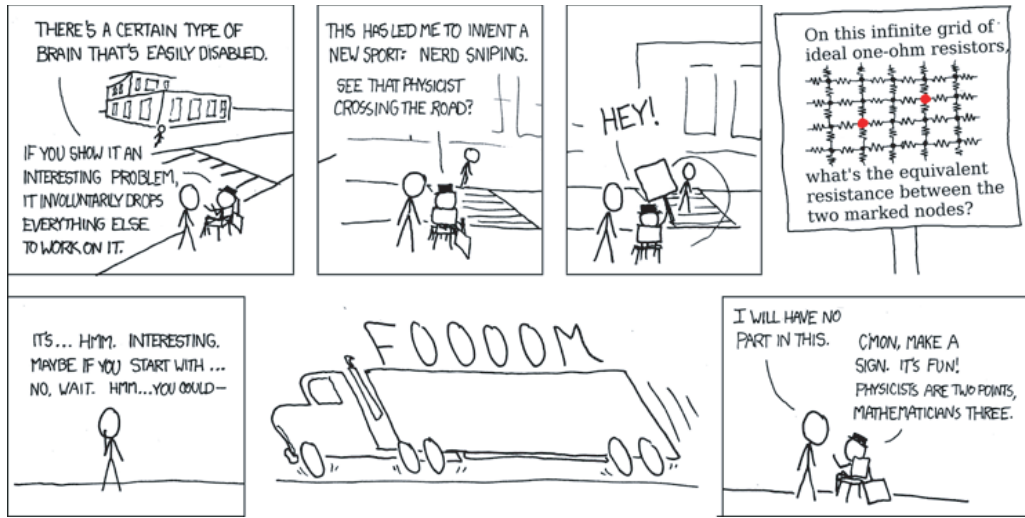of the mysteries of modern science.

FIGURE 32 An electrical problem that is not easy (© Randall Munroe).
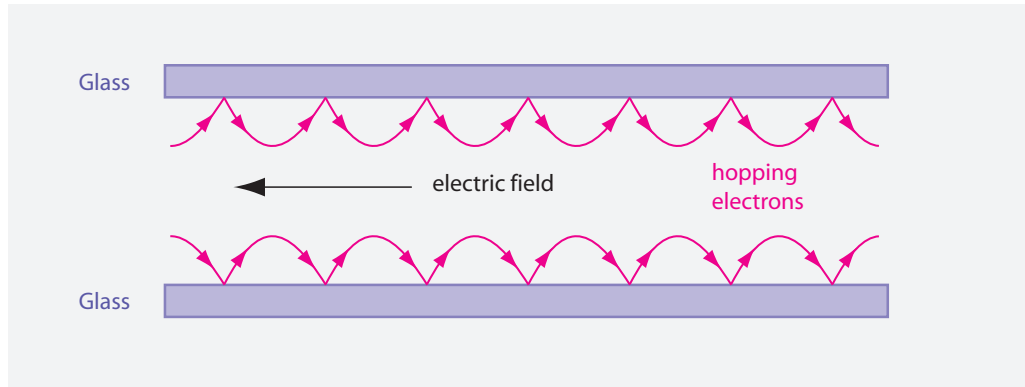

FIGURE 33 Free electrons can hop along a glass wall.

* *

A pure magnetic field cannot be transformed into a pure electric field by change of observation frame. The best that can be achieved is a state similar to an equal mixture of magnetic and electric fields. Can you provide an argument elucidating this relation?

Challenge 64 ny

* *

Calculating resistance of infinite grids is one of the most captivating problems in electricity, as shown in Figure 32. Can you find the solution?

Challenge 65 ny

HOPPING ELECTRONS AND THE BIGGEST DISAPPOINTMENT OF THE TELEVISION INDUSTRY

It is well known that when an electric field in a vacuum points along a glass surface, electrons can *hop* along the glass surface. The general effect is shown in Figure 33; usually, the effect is unwelcome. Among others, the hopping effect is responsible for sparks in

vacuum systems that contain high voltage.

When this effect was studied in detail, it turned out that reasonably low electric fields are sufficient to create sizeable electric hopping currents. The effect also works around bends and corners. Furthermore, electric switches that change the hopping direction can be constructed. In short, the hopping effect can be used to make extremely cheap flat television displays of high image quality. The idea is to put an array of electron sources – essentially sharp metal tips – at the start of glass channels and to transport the emitted electrons along the channels, making use of suitable switches, until they hit phosphores-cent colour pixels. These are the same pixels that were used in the then common – bulky and heavy – television tubes. Since the hopping effect also works around bends and cor-ners, and since it only needs glass and a bit of metal, the whole system can be made ex-tremely flat, lightweight and cheap. Already in the early 1990s, the laboratory samples of the electron hopping displays were spectacularly good: the small displays were brighter, sharper and cheaper than liquid crystal displays, and the large ones brighter, sharper and cheaper than plasma displays. Affordable flat television was on the horizon.

Then came the disappointment. The lifetime of the displays was only of the order of one hundred hours. Despite the most intense material research possible, achieving a higher lifetime turned out to be impossible. All tricks that were tried did not help. De-spite all their fantastic properties, despite huge investments in the technology, despite the best material researchers working on the issue, electron hopping displays could not be brought to market. Not a single display was ever sold.

How do nerves work?

Nerves are wonders. Without nerves, we would not experience pleasure, we would not experience pain, we would not see, we would not hear. Without nerves, we would not live. But how do nerves transport signals?

In 1789, as mentioned above, Luigi Galvani discovered that nerves transport electric signals,0 by doing experiments with frog legs. Are nerves wires? One and a half centuries after Galvani it became clear that nerves do not conduct electricity using electrons, as metal wires do, but by using *ions*. Nerve signals propagate using the motion of sodium Na^+ and potassium K^+ ions through the cell membrane of the nerve. The resulting signal speed is between 0.5 m/s and 120 m/s, depending on the type of nerve. (Nerve axons coated with myelin, a protein that acts as an electric insulator, are faster than uncoated axons.) The signal speed is sufficient for the survival of most species – it helps the body to run away in case of danger.

Nerves differ from wires in another aspect: they cannot transmit constant voltage sig-nals, but only signal *pulses*. The first, approximate model for this behaviour was presented in 1952 by Hodgkin and Huxley. Using observations about the behaviour of potassium and sodium ions, they deduced an elaborate evolution equation that describes the volt-age V in nerves, and thus the way the signals propagate. The equation reproduces the characteristic voltage spikes measured in nerves, shown in Figure 34.

The precise mechanism with which ions cross the membranes, using so-called *channel proteins*, was elucidated only twenty years later. Despite this huge body of work, and even though Hodgkin and Huxley received the Nobel Prize for Medicine for their work, their model cannot be correct. The model does not explain several observation about nerves: it
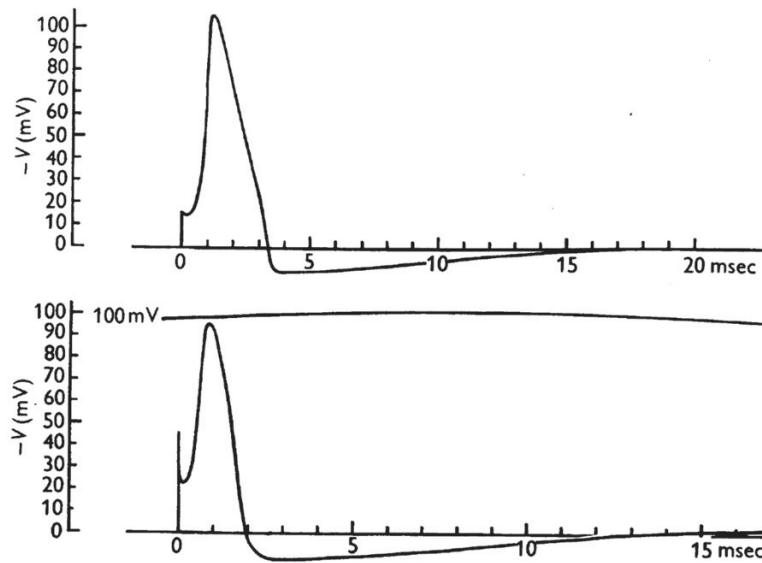
Fig. 13. Upper curve: solution of eqn. (26) for initial depolarization of 15 mV, calculated for 6° C. Lower curve: tracing of membrane action potential recorded at 9·1° C (axon 14). The vertical scales are the same in both curves (apart from curvature in the lower record). The horizontal scales differ by a factor appropriate to the temperature difference.

FIGURE 34 The electrical signals calculated (above) and measured (below) in a nerve, following Hodgkin and Huxley.

does not explain the reversibility of the propagation process, the thickness change of the nerve during propagation, the excitation of nerves by simple deformation or temperature changes; most of all, the does not explain the working of anesthetics. The working of nerves remained unknown.

Ref. 31   Only around the year 2000 did Thomas Heimburg and his team discover the way signals propagate in nerves. They showed that a nerve pulse is an electromechanical solitonic wave of the cylindrical membrane. In the cylindrical membrane, the protein structure changes from liquid to solid and back to liquid. A short, slightly thicker ring of solid proteins propagates along the cylinder: that is the nerve pulse. The model is shown in Figure 35. (The term 'solid' has a precise technical meaning in two-dimensional systems and describes a specific ordered state of the molecules.) This model explains all the properties of nerve pulses that were unexplained before. In particular, it explains that anaesthetics work because they dissolve in the membrane and thus block the formation and the propagation of the rings. All quantitative predictions of the model match observations. In short, nerve signals are electromechanical pulses; they are a mixture of current and sound waves.

A SUMMARY: THREE BASIC FACTS ABOUT ELECTRICITY

The experiments we have described so far show three basic results:

▷ *Electric charges exert forces on other charges.*
▷ *Electric charges are conserved.*
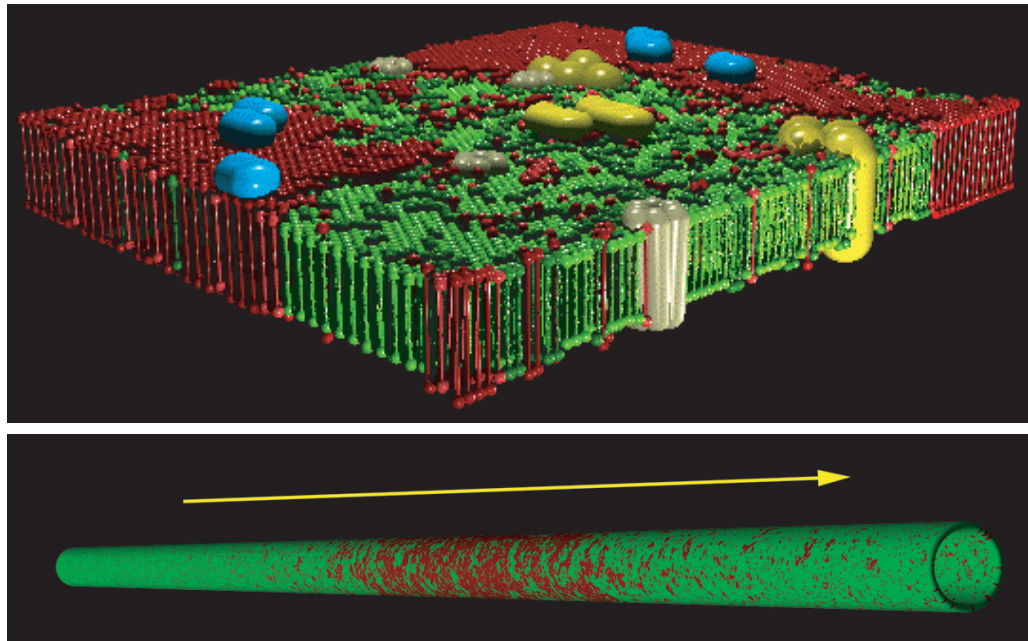
FIGURE 35 Top: A biomembrane, with solid-ordered lipids (red), liquid lipids (green) and various dissolved proteins (yellow, blue, white). Bottom: a nerve pulse propagating as a two-dimensional phase transformation liquid/solid/liquid along a cylindrical nerve membrane (© Thomas Heimburg/Wiley-VCH).

⊳ *Charges, like all matter, move slower than light.*

From these three statements – the definition of charge, the conservation of charge, and the invariance of the speed of light – we can deduce every aspect of classical electrodynamics. (If we want, we can add the non-existence of magnetic charge as an explicit, additional assumption.) In particular, the Lagrangian of electrodynamics and Maxwell's field equations can be deduced from these three statements; they describe the way that charges *produce* any electric, magnetic or electromagnetic field. Also the Lorentz force can be deduced; it describes how the motion of massive charges and the motion of the electromagnetic field is related.

The proof of the connection between charge conservation and the field equations can
Ref. 32   be given mathematically; we do not present it here, because the algebra is somewhat involved. The essential connection is: all of electrodynamics follows from the properties
Ref. 33   of charges that we have discovered so far.
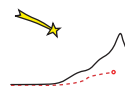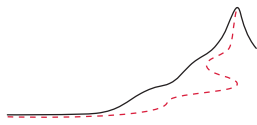
THE DESCRIPTION OF ELECTROMAGNETIC FIELD EVOLUTION

LECTRIC and magnetic fields change: simply said, they move. How exactly does this happen? In the 1860s, James Clerk Maxwell* collected all xperimental knowledge he could find, and deduced the precise description of electromagnetic field motion. Twenty years later, Heaviside and Hertz extracted the main points of Maxwell ideas, calling their summary *Maxwell's theory of the electromagnetic field*.

The motion of the electromagnetic field is described by a set of evolution equations. In the relativistic description, the set consists of *two* equations, in the non-relativistic case of *four* equations. All observations of classical electrodynamics follow from these equations. In fact, if quantum effects are properly taken into account, *all* electromagnetic effects of nature are described.

THE FIRST FIELD EQUATION OF ELECTRODYNAMICS

The first relativistic field equation of electrodynamics is the precise statement that electromagnetic fields *originate at charges*, and nowhere else. It can variously be written**

$$d\boldsymbol{F} = j\mu_0 \quad \text{or}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{E} = \frac{\rho}{\varepsilon_0} \quad \text{and} \quad \boldsymbol{\nabla}\times\boldsymbol{B} - \frac{1}{c^2}\frac{\partial\boldsymbol{E}}{\partial t} = \mu_0\boldsymbol{j} \ . \tag{26}$$

---

* James Clerk Maxwell (b. 1831 Edinburgh, d. 1879 Cambridge), Scottish physicist. He founded electromagnetism by theoretically unifying electricity and magnetism, as described in this chapter. His work on thermodynamics forms the second pillar of his activity. In addition, he studied the theory of colours and developed the colour triangle; he was one of the first people to make a colour photograph. He is regarded by many as the greatest physicist ever. Both 'Clerk' and 'Maxwell' were his family names.

** There is a certain freedom in writing the equations, because different authors absorb different combinations of the constants c and μ_0 into the definitions of the quantities F, A and j. This is the most common version. The equations can be generalized to cases where the charges are not surrounded by vacuum, but located inside matter. We will not explore these situations in our walk because, as we will see during our mountain ascent, the seemingly special case of vacuum in fact describes all of nature.
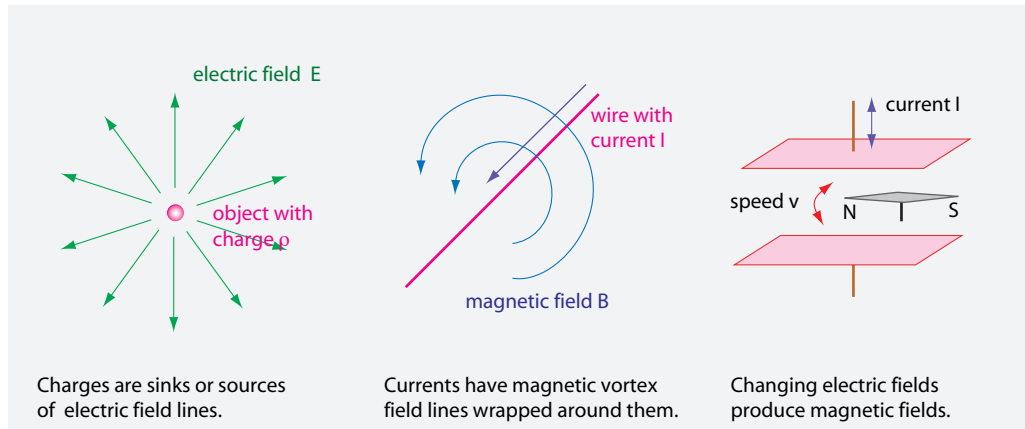
electric field E

object with charge ρ

wire with current I

magnetic field B

current I

speed v   N | S

Charges are sinks or sources of electric field lines.

Currents have magnetic vortex field lines wrapped around them.

Changing electric fields produce magnetic fields.

FIGURE 36 The first of Maxwell's field equations of electrodynamics illustrated in three drawings.

Each of these two equivalent ways[*] to write the first Maxwell equation makes a simple statement: *electrical charges carry the electromagnetic field*; they carry it along with them. The first equation thus describes the attraction of dust by electrically charged objects and the working of electromagnets.

This first equation is equivalent to the three basic observations illustrated in Figure 36: Coulomb's relation, Ampère's relation, and the way changing electrical fields induce magnetic effects. More precisely, if we know where charges are and how they move, we can determine the electromagnetic field F they generate. Static charges, described by a density ρ, produce electrostatic fields, and moving charges, described by a 3-current density j, produce a mix of electric and magnetic fields. Stationary currents produce magneto-static fields.

Challenge 66 e

The first equation also contains the *right hand rule* for magnetic fields around wires, through the vector product. As mentioned, the equation also states, most clearly in its last form, that changing electric fields induce magnetic fields. The effect is essential in the primary side of transformers. The small factor $1/c^2$ implies that the effect is small; therefore coils with *many* windings or *strong* electric currents are needed to produce or detect the effect.

THE SECOND FIELD EQUATION OF ELECTRODYNAMICS

The second of Maxwell's equations, illustrated in Figure 37, expresses the observation that in nature there are no magnetic charges, i.e., that magnetic fields have no sources. As a result, the equation also gives a precise description of how changing magnetic fields cre-

---

[*] In component form, the first equation can be written

$$d_\mu F^{\mu\nu} = j^\nu \mu_0 = (\rho c, j)\mu_0 = (\rho_0 \gamma c, \rho_0 \gamma v)\mu_0 \quad \text{or}$$

$$(\partial_t/c, \partial_x, \partial_y, \partial_z) \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix} = \mu_0(\rho c, j) \, . \tag{27}$$
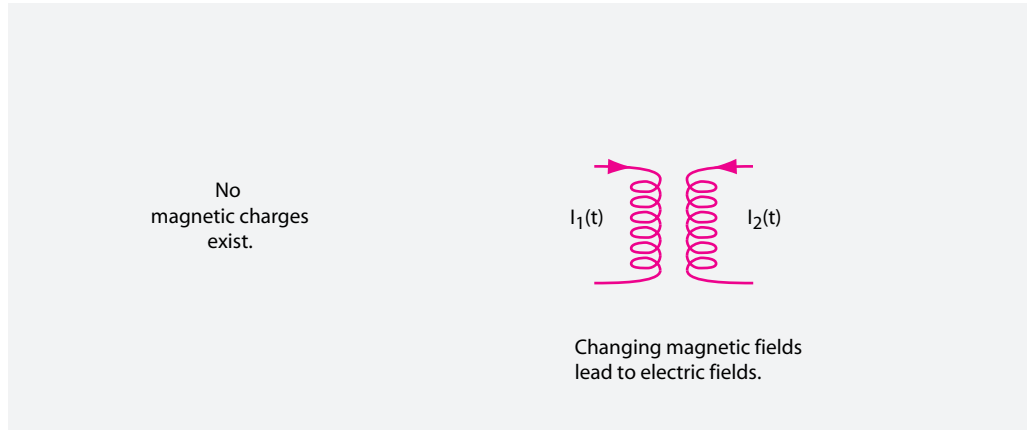
FIGURE 37 The second field equation of electrodynamics.

ate electric fields, and vice versa. The second of Maxwell's equations for electrodynamics can variously be written

$$d\ {}^{*}\boldsymbol{F} = 0 \quad \text{with} \quad {}^{*}\boldsymbol{F}^{\rho\sigma} = \frac{1}{2}\varepsilon^{\rho\sigma\mu\nu}\boldsymbol{F}_{\mu\nu} \quad \text{or}$$

$$\nabla{\cdot}\boldsymbol{B} = 0 \quad \text{and} \quad \nabla{\times}\boldsymbol{E} = -\frac{\partial\boldsymbol{B}}{\partial t}\ . \tag{28}$$

First of all, the second field equation[*] thus expresses the *lack of sources for the dual field tensor* ${}^{*}\boldsymbol{F}$. In other words, in nature there are no magnetic charges, i.e., no magnetic monopoles: there are no sources for magnetic fields. The equation thus states that cutting a magnet with a north and a south pole in any way always produces pieces with *two* poles, never a piece with a single pole.

Since there are no magnetic charges, magnetic field lines are *always closed*; not only the magnetic field lines induced by charges, no, *all* magnetic field lines are vortex lines. For example, field lines continue inside magnets. This is often expressed mathematically by stating that the *magnetic flux* through a *closed* surface S – such as a sphere or a cube – *always vanishes*: $\int_{S}\boldsymbol{B}\,\mathrm{d}\boldsymbol{A} = 0$. In other words, all field lines that enter a closed volume also leave it.

---

[*] In component form, the second equation can be written

$$d_{\mu}{}^{*}\boldsymbol{F}^{\mu\nu} = 0 \quad \text{or}$$

$$(\partial_t/c, \partial_x, \partial_y, \partial_z)\begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z/c & -E_y/c \\ B_y & -E_z/c & 0 & E_x/c \\ B_z & E_y/c & -E_x/c & 0 \end{pmatrix} = (0, 0, 0, 0) \quad \text{or}$$

$$\varepsilon^{\sigma\mu\nu\rho}\partial_{\mu}\boldsymbol{F}_{\nu\rho} = 0 \quad \text{or}$$

$$\partial_{\mu}\boldsymbol{F}_{\nu\rho} + \partial_{\nu}\boldsymbol{F}_{\rho\mu} + \partial_{\rho}\boldsymbol{F}_{\mu\nu} = 0 \quad , \text{or} \quad . \tag{29}$$

We note that the dual tensor ${}^{*}\boldsymbol{F}$ follows form the field tensor $\boldsymbol{F}$ by substituting E/c by B and B by $-E/c$. This is the so-called *duality transformation*. More on this duality below.

Furthermore, the second Maxwell equation expresses that changes in magnetic fields produce electric fields: this effect is used in the secondary side of transformers and in dynamos. The cross product in the expression implies that an electric field generated in this way – also called an *electromotive field* – has no start and end points. The electromotive field lines thus run in circles: in most practical cases they run along electric circuits. In short, an electric field can (also) have vortices, but only when there is a changing magnetic field. The minus sign is essential to ensure energy conservation (why?) and has a special name: it is called *Lenz's rule*.

In practice, the second Maxwell equation is always needed together with the first. Can you see why?

The validity and the essence of Maxwell's field equations

Together with Lorentz' evolution equation

$$m\boldsymbol{b} = q\boldsymbol{F}u \quad \text{or}$$
$$\mathrm{d}E/\mathrm{d}t = q\boldsymbol{E}\boldsymbol{v} \quad \text{and} \quad \mathrm{d}\boldsymbol{p}/\mathrm{d}t = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \ . \tag{30}$$

which describes how charges move given the motion of the fields, Maxwell's evolution equations (26) and (28) describe *all* electromagnetic phenomena occurring on everyday scales, from mobile phones, car batteries, to personal computers, lasers, lightning, holograms and rainbows. This description of electromagnetic fields is complete for everyday life. Only quantum effects and the effects of curved space-time are not included.

Maxwell's equations seem very complex. But we should never forget that they contain only *four* basic ideas. First: electric charges follow Coulomb's rule. Second: electric charges moves slower than light. Third: electric charges are conserved. Fourth: magnetic charges do not exist. If we want to be very simplistic, Maxwell's equations are just the relativistic formulation of Coulomb's rule.

We will not study many applications of the field equations. True, the range of applications is vast: modern medicine, transport, telecommunication, computers, and most jobs and many pleasures depend on electricity. But we leave these topics aside and continue directly towards our aim to understand the connection between electromagnetic fields, everyday motion and the motion of light. In fact, the electromagnetic field has an important property that we mentioned right at the start: the field itself can move. In particular, the field can carry energy, linear momentum and angular momentum.

Colliding charged particles

Electromagnetic fields move. A simple experiment clarifies the meaning of motion for fields: When two charged particles collide, their total momentum is *not* conserved. Let us check this.

Imagine two particles of identical mass and identical charge just after a collision, when they are moving away from one another. The situation is illustrated in Figure 38. Imagine also that the two masses are large, so that the acceleration due to their electrical repulsion is small. For an observer at the centre of gravity of the two, each particle feels an acceleration from the electric field of the other. This electric field E is given by the so-called
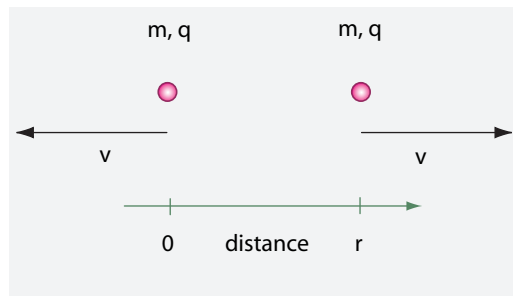
FIGURE 38 Charged particles after a collision.

Heaviside formula

$$E = \frac{q\,(1 - v^2/c^2)}{4\pi e_0 r^2}\,.\qquad(31)$$

In other words, the total system has a vanishing total momentum for this observer.

Take a second observer, moving with respect to the first with velocity v, so that the first charge will be at rest. Expression (31) leads to two *different* values for the electric fields, one at the position of each particle. In other words, the system of the two particles is not in inertial motion, as we would expect; the total momentum is not conserved for this observer. The missing momentum is small, but where did it go?

This at first surprising effect has even been put in the form of a theorem by Van Dam and Wigner. They showed that, for a system of particles interacting at a distance, the total particle energy–momentum cannot remain constant in all inertial frames.

The total momentum of the system is conserved only because the electromagnetic field itself also carries some momentum. In short, momentum is conserved in the experiment, but some of it is carried by the field. The precise amount depends on the observer.

Two colliding charged particles thus show us that electromagnetic fields have momentum. If electromagnetic fields have momentum, they are able to *strike* objects and to be struck by them. As we will show below, light is also an electromagnetic field. Thus we should be able to move objects by shining light on to them. We should even be able to suspend particles in mid air by shining light on to them from below. Both predictions are correct, and some experiments will be presented shortly.

We conclude that any sort of field leading to particle interactions must carry both energy and momentum, as the argument applies to all such cases. In particular, it applies to nuclear interactions. Indeed, in the quantum part of our mountain ascent we will even find an additional result: all fields are themselves composed of particles. The energy and momentum of fields then become an obvious state of affairs. In short, it makes sense to say that electromagnetic fields move, because they carry energy and momentum.

THE GAUGE FIELD – THE ELECTROMAGNETIC VECTOR POTENTIAL

The study of moving fields is called *field theory* and electrodynamics is the prime example. (The other classical example is fluid dynamics; moving electromagnetic fields and moving fluids are very similar mathematically.) Field theory is a beautiful topic; field lines, equipotential lines and vortex lines are some of the concepts introduced in this domain.
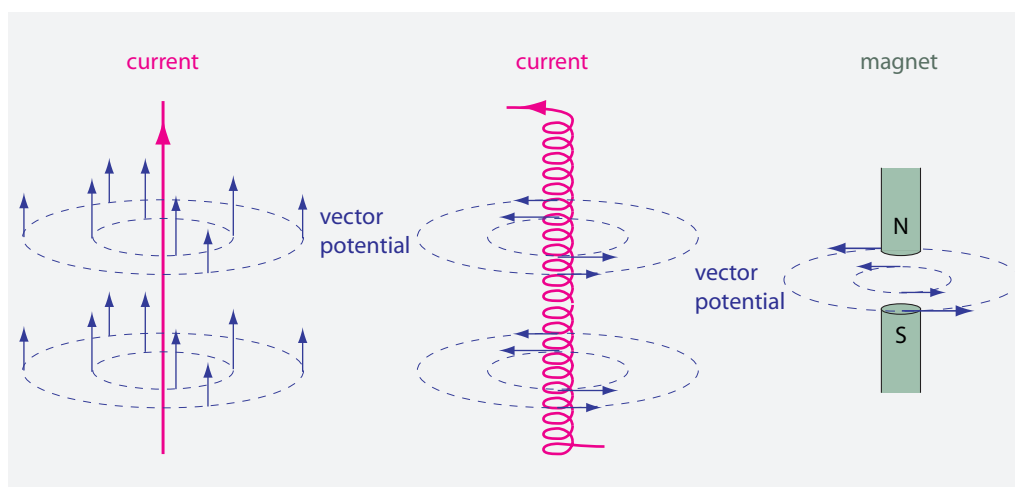
FIGURE 39 Vector potentials for selected situations.

They fascinate many.* However, in this mountain ascent we keep the discussion focused on motion.

We have seen that fields force us to extend our concept of motion. Motion is not only the change in state of objects and of space-time, but also the *change in state of fields*. We therefore need, also for fields, a complete and precise description of their state.

The observations using amber and magnets have shown us that *electromagnetic fields possess energy and momentum*. Fields can impart energy and momentum to particles. The experiments with motors have shown us that objects can add energy and momentum to fields. We therefore need to define a *state function* which allows us to define energy and momentum for electric and magnetic fields. And since electric and magnetic fields transport energy, their motion must follow the speed limit in nature.

Hertz and Heaviside defined the state function of fields in two standard steps. The first step is the definition of the *(magnetic) vector potential*, which describes the momentum per charge that the field provides:

$$A = \frac{p}{q} \ . \tag{32}$$

When a charged particle moves through a magnetic potential $A(x)$, its momentum changes by $q\Delta A$; it changes by the difference between the potential values at the start and end points, multiplied by its charge. Owing to this definition, the vector potential has the property that

$$B = \nabla \times A = \text{curl } A \tag{33}$$

i.e., that the magnetic field is the *curl* of the magnetic potential. In most other languages the curl is called the *rotation* and abbreviated rot. To visualize what the curl or rotation is,

Ref. 36

Challenge 71 s

Ref. 1, Ref. 19

---

* What is the relation, for static fields, between field lines and (equi-) potential surfaces? Can a field line cross a potential surface twice? For more details on topics such as these, see the *free* textbook by Bo Thidé, *Electromagnetic Field Theory*, on his www.plasma.uu.se/CED/Book website. And of course, in English, have a look at the texts by Schwinger and by Jackson.
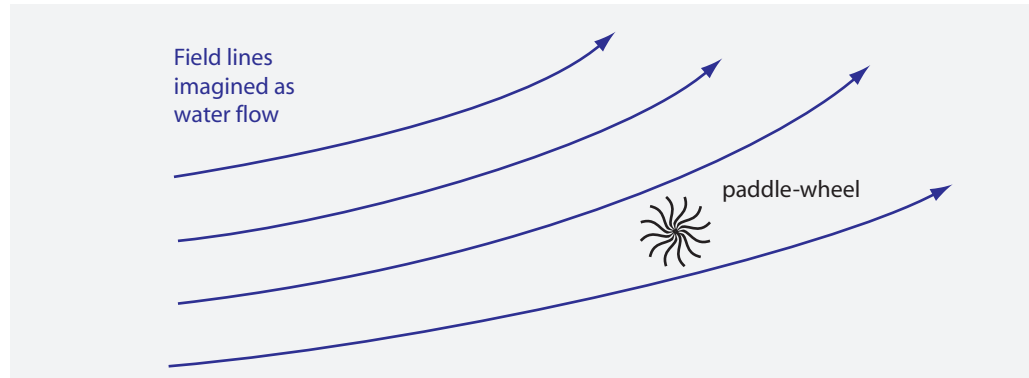
FIGURE 40 Visualizing the curl of a vector field. Imagine the field to be flowing air and check whether the small paddle-wheel rotates; if it does, the local curl is non-zero. The direction of the curl is the direction of the paddle-wheel axis that yields the highest rotation velocity.

imagine that the field vectors are the velocity vectors of flowing air. Now put a tiny paddle-wheel at a point, as shown in Figure 40. If it turns, the curl is non-zero. The rotation speed of the paddle-wheel is maximal for some direction of the axis; this maximal speed defines both the magnitude and the direction of the curl at the point. (The right-hand rule is implied.) For example, the curl for the velocities of a rotating solid body is everywhere 2ω, or twice the angular velocity.

The vector potential for a long straight current-carrying wire is parallel to the wire; it has the magnitude

$$A(r) = -\frac{\mu_0 I}{4\pi} \ln \frac{r}{r_0} \, , \tag{34}$$

which depends on the radial distance r from the wire and an integration constant r_0. This expression for the vector potential, pictured in Figure 39, shows how the moving current produces a linear momentum in the (electro-) magnetic field around it. In the case of a solenoid, the vector potential 'circulates' around the solenoid. The magnitude obeys

$$A(r) = -\frac{\Phi}{4\pi}\frac{1}{r} \, , \tag{35}$$

where Φ is the magnetic flux inside the solenoid. We see that, in general, the vector potential is *dragged along* by moving charges. The dragging effect decreases for larger distances. This fits well with the image of the vector potential as the momentum of the electromagnetic field.

This behaviour of the vector potential around charges is reminiscent of the way honey is dragged along by a spoon moving in it. In both cases, the dragging effect decreases with distance. However, the vector potential, unlike the honey, does *not* produce any friction that slows down charge motion. The vector potential thus behaves like a frictionless liquid.

Inside the solenoid, the magnetic field is constant and uniform. For such a field **B** we

find the vector potential

$$A(\mathrm{r}) = -\frac{1}{2}\boldsymbol{B} \times \boldsymbol{r} \ . \tag{36}$$

In this case, the magnetic potential thus increases with increasing distance from the origin.* In the centre of the solenoid, the potential vanishes. The analogy of the dragged honey gives exactly the same behaviour.

However, there is a catch. The magnetic potential is *not* defined uniquely. If $\boldsymbol{A}(\boldsymbol{x})$ is a vector potential, then the different vector potential

$$\boldsymbol{A}'(\boldsymbol{x}) = \boldsymbol{A}(\boldsymbol{x}) + \boldsymbol{\nabla}\Lambda \ , \tag{37}$$

where $\Lambda(t, \boldsymbol{x})$ is some scalar function, is *also* a vector potential for the same situation. (The magnetic field $\boldsymbol{B}$ stays the same, though.) Worse, can you confirm that the corre-
sponding (absolute) momentum values also change? This unavoidable ambiguity, called *gauge invariance* or *gauge symmetry*, is a central property of the electromagnetic field. We will explore it in more detail below.

Not only the momentum, but also the energy of the electromagnetic field is defined ambiguously. Indeed, the second step in the specification of a state for the electromag-
netic field is the definition of the *electric potential* as the energy U per charge:

$$\varphi = \frac{U}{q} \tag{38}$$

In other words, the potential $\varphi(\boldsymbol{x})$ at a point $\boldsymbol{x}$ is the energy needed to move a unit charge to the point $\boldsymbol{x}$ starting from a point where the potential vanishes. The potential energy is thus given by $q\varphi$. From this definition, the electric field $\boldsymbol{E}$ is simply the *change* of the potential with position corrected by the time dependence of momentum, i.e.,

$$\boldsymbol{E} = -\boldsymbol{\nabla}\varphi - \frac{\partial}{\partial t}\boldsymbol{A} \ , \tag{39}$$

Obviously, there is a freedom in the choice of the definition of the potential. If $\varphi(\boldsymbol{x})$ is a possible potential, then

$$\varphi'(\boldsymbol{x}) = \varphi(\boldsymbol{x}) - \frac{\partial}{\partial t}\Lambda \tag{40}$$

is also a potential function for the same situation. This freedom is the generalization of the freedom to define energy up to a constant. Nevertheless, the electric field $\boldsymbol{E}$ remains the same for all potentials.
To be convinced that the potentials really are the energy and momentum of the elec-

---

* This is only possible as long as the field is constant; since all fields drop again at large distances – because the energy of a field is always finite – also the vector potential drops at large distances.

tromagnetic field, we note that for a moving charge we have

$$\frac{d}{dt}\left(\frac{1}{2}mv^2 + q\varphi\right) = \frac{\partial}{\partial t}q\ (\varphi - \boldsymbol{v}\boldsymbol{A})$$
$$\frac{d}{dt}\left(m\boldsymbol{v} + q\boldsymbol{A}\right) = -\boldsymbol{\nabla}q\ (\varphi - \boldsymbol{v}\boldsymbol{A})\ , \tag{41}$$

which show that the changes of generalized energy and momentum of a particle (on the left-hand side) are due to the change of the energy and momentum of the electromagnetic field (on the right-hand side).*

In relativistic 4-vector notation, the energy and the momentum of the field appear together in one quantity. The state function of the electromagnetic field becomes

$$A^\mu = (\varphi/c, \boldsymbol{A}) \tag{42}$$

and is called the *4-potential*. It is easy to see that the description of the field is complete, since we have

$$\boldsymbol{F} = d\,A \quad \text{or} \quad \boldsymbol{F}^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \quad (\text{and} \quad \boldsymbol{F}_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu)\ , \tag{43}$$

which means that the electromagnetic field F is completely specified by the 4-potential A.** But as just said, the 4-potential itself is *not* uniquely defined. Indeed, any other equivalent 4-potential A' is related to A by the *gauge transformation*

$$A'^{\,\mu} = A^\mu + \partial^\mu \Lambda \tag{44}$$

where $\Lambda = \Lambda(t, x)$ is any arbitrarily chosen scalar field. The new field A' leads to the *same* electromagnetic field, and to the same accelerations and evolutions. The 4-potential A is thus an *overdescription* of the physical situation as several *different* gauge choices correspond to the *same* physical situation.*** Therefore we have to check that all measurement results are independent of gauge transformations, i.e., that all observables are gauge invariant quantities. Such gauge invariant quantities are, as we just saw, the fields $\boldsymbol{F}$ and $^*\boldsymbol{F}$, and in general all classical quantities. We add that many theoretical physicists use the term 'electromagnetic field' loosely for both the quantities $F^{\mu\nu}$ and A^μ.

There is a simple image, due to Maxwell, to help overcoming the conceptual difficulties of the vector potential. It turns out that the closed line integral over A_μ is gauge invariant, because

$$\oint A^\mu dx_\mu = \oint (A^\mu + \partial^\mu \Lambda)dx_\mu = \oint A'^{\,\mu} dx_\mu\ . \tag{45}$$

---

* This connection also shows why the expression $P^\mu - qA^\mu$ appears so regularly in formulae; indeed, it plays a central role in the quantum theory of a particle in the electromagnetic field.

** The connection between A_μ and A^μ, the same as for every other 4-vector, was mentioned earlier on; can you restate it?

*** Choosing a function Λ is often called *choosing a gauge*; the 4-potential A is also called the *gauge field*. These strange terms have historic reasons and are now common to all of physics.

In other words, if we picture the vector potential as a quantity allowing us to associate a number to a tiny ring at each point in space, we get a good, gauge invariant picture of the vector potential.*

Now that we have defined a state function that describes the energy and momentum of the electromagnetic field, let us look at what happens in more detail when electromagnetic fields move.

Energy and momenta of the electromagnetic field

All moving entities have energy, momentum and angular momentum. This also applies to the electromagnetic field. Indeed, the description so far allows us to write the *total* energy E_{nergy} of the electromagnetic field as

$$E_{\text{nergy}} = \frac{1}{4\pi} \int \frac{\varepsilon_0}{2}(\boldsymbol{E}^2 + c^2\boldsymbol{B}^2) \, \mathrm{d}V \; . \tag{46}$$

Energy is thus quadratic in the fields.

For the total linear momentum $\boldsymbol{p}$ we obtain

$$\boldsymbol{p} = \frac{1}{4\pi} \int \varepsilon_0 \boldsymbol{E} \times \boldsymbol{B} \, \mathrm{d}V \; . \tag{47}$$

The expression inside the integral, is the *momentum density*. The vector $\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{B}/\mu_0$, is called the *Poynting vector*.** The Poynting vector $\boldsymbol{S}$ is the *energy flux*, and has the units W/m^2. It is the momentum density divided by c^2; indeed, special relativity implies that the momentum and the energy flow for electromagnetic fields are related by a factor c^2. The Poynting vector thus describes the energy flowing per area per time, in other words, the power per area. As shown below, the Poynting vector is a part of the energy–momentum tensor.

For the total angular momentum we have

$$\boldsymbol{L} = \frac{\varepsilon_0}{4\pi} \int \boldsymbol{E} \times \boldsymbol{A} \, \mathrm{d}V = \frac{\varepsilon_0}{4\pi} \int \boldsymbol{r} \times (\boldsymbol{E} \times \boldsymbol{B}) \, \mathrm{d}V \; , \tag{48}$$

where $\boldsymbol{A}$ is the magnetic vector potential.

In summary, the electromagnetic field has energy and momenta. Nevertheless, for most everyday situations, the values are negligibly small, as you may want to check.

The Lagrangian of electromagnetism

The motion of a charged particle and the related motion of the electromagnetic field can also be described using a Lagrangian, instead of using the three equations given above. It is not hard to see that the action S_{CED} for a particle in classical electrodynamics can be

Challenge 79 e

Ref. 38

---

* In the part of the text on quantum theory we will see that the exponent of this expression, namely $\exp(iq \oint A_\mu \mathrm{d}x^\mu)/\hbar$, usually called the *phase factor*, can indeed be directly observed in experiments.
** John Henry Poynting (1852–1914) introduced the concept in 1884.

symbolically defined by[*]

$$S_{\text{CED}} = -mc^2 \int \mathrm{d}\tau - \frac{1}{4\mu_0} \int \boldsymbol{F} \wedge * \boldsymbol{F} - \int j \wedge A \, , \tag{49}$$

which in index notation becomes

$$S_{\text{CED}} = -mc \int_{-\infty}^{\infty} \sqrt{\eta_{\mu\nu} \frac{\mathrm{d}x_n^\mu(s)}{\mathrm{d}s} \frac{\mathrm{d}x_n^\nu(s)}{\mathrm{d}s}} \, \mathrm{d}s - \int_{\mathbf{M}} \left(\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} + j_\mu A^\mu \right) \mathrm{d}^4 x \, , \tag{50}$$

or, in 3-vector notation

$$S_{\text{CED}} = -mc^2 \int \mathrm{d}\tau + \int (qvA - q\varphi) \, \mathrm{d}t \mathrm{d}V + \int \left(\frac{\varepsilon_0}{2} E^2 - \frac{1}{2\mu_0} B^2 \right) \mathrm{d}t \mathrm{d}V \, . \tag{51}$$

The new part is the measure of the change – or action – due to the electromagnetic field. The pure field change is given by the term $\boldsymbol{F} \wedge * \boldsymbol{F}$, and the change due to interaction with matter is given by the term $j \wedge A$.

The least action principle, as usual, states that the change in a system is always as small as possible. The action S_{CED} leads to the evolution equations by requiring that the action be stationary under variations δ and δ' of the positions and of the fields which vanish at infinity. In other terms, the principle of least action requires that

$$
\begin{aligned}
\delta S = 0 \quad &\text{when} \quad x_\mu = x_\mu + \delta_\mu \quad \text{and} \quad A_\mu = A_\mu + \delta'_\mu \quad , \\
&\text{provided} \quad \delta x_\mu(\theta) \to 0 \quad \text{for} \quad |\theta| \to \infty \\
&\text{and} \quad \delta A_\mu(x_\nu) \to 0 \quad \text{for} \quad |x_\nu| \to \infty \, .
\end{aligned}
\tag{52}
$$

In the same way as in the case of mechanics, using the variational method for the two variables A and x, we recover the evolution equations for particle position and fields

$$b^\mu = \frac{q}{m} F_\nu^\mu \, u^\nu \quad , \quad \partial_\mu F^{\mu\nu} = j^\nu \mu_0 \quad , \quad \text{and} \quad \varepsilon^{\mu\nu\rho\sigma} \partial_\nu F_{\rho\sigma} = 0 \, , \tag{53}$$

which we know already: they are the Lorentz relation and the two field equations. Obviously, they are equivalent to the variational principle based on S_{CED}. Both descriptions have to be completed by specifying *initial conditions* for the particles and the fields, as well as *boundary conditions* for the latter. We need the first and zeroth derivatives of the position of the particles, and the zeroth derivative for the electromagnetic field.

With the Lagrangian (49) all of classical electrodynamics can be described and understood. For the rest of our exploration of electrodynamics, we look at some specific topics from this vast field.

---

[*] The product described by the symbol $\wedge$, 'wedge' or 'hat', and the duality operator $*$ have a precise mathematical meaning. The background, the concept of *(mathematical) form*, carries us too far from our walk.

THE ENERGY–MOMENTUM TENSOR AND ITS SYMMETRIES OF MOTION

We know from classical mechanics that we get the definition of energy and momentum by using Noether's theorem. In particular, both the definition and the conservation of energy and momentum arise from the Lorentz symmetry of the Lagrangian. For example, we found that relativistic particles have an energy–momentum *vector*. At the point at which the particle is located, it describes its energy and momentum.

Since the electromagnetic field is not a localized entity, like a point particle, but an extended entity, a full description is more involved. In order to describe the energy–momentum of the electromagnetic field completely, we need to know the *flow* of energy and momentum at every point in space, separately *for each direction*. This makes a description with a *tensor* necessary, the so-called *energy–momentum tensor* T of the electromagnetic field.

The electric field times a charge is the force on that charge, or equivalently, its momentum increase per time. The generalization for the full electromagnetic field F, and for the full power–force (or 4-force) vector K is

$$F^{\mu\nu} j_\mu = K^\nu = \partial_\mu T^{\mu\nu} \quad . \tag{54}$$

This short equation, which can also be derived from the Lagrangian, contains a lot of information. In particular, it expresses that every change in energy of the field is the sum of the energy radiated away (via the energy flow described by the Poynting vector S) and of change in the kinetic energy of the charges. The equation also makes a similar statement on the momentum of the electromagnetic field.

The detailed parts of the energy–momentum tensor T are found to be

$$T^{\mu\nu} = \begin{pmatrix} \begin{array}{c} \text{energy} \\ \text{density} \end{array} & \begin{array}{c} \text{energy flow or} \\ \text{momentum density} \end{array} \\ \begin{array}{c} \text{energy flow or} \\ \text{momentum density} \end{array} & \begin{array}{c} \text{momentum} \\ \text{flow density} \end{array} \end{pmatrix}$$

$$= \left(\begin{array}{c|c} u & S/c = c\boldsymbol{p} \\ \hline c\boldsymbol{p} & T \end{array} \right) = \left(\begin{array}{c|c} (\varepsilon_0 E^2 + B^2/\mu_0)/2 & \varepsilon_0 c \boldsymbol{E} \times \boldsymbol{B} \\ \hline \begin{array}{c} \varepsilon_0 c \cdot \\ \boldsymbol{E} \times \boldsymbol{B} \end{array} & \begin{array}{c} -\varepsilon_0 E_i E_j - B_i B_j/\mu_0 \\ 1/2 \delta_{ij}(\varepsilon_0 E^2 + B^2/\mu_0) \end{array} \end{array} \right) \tag{55}$$

where $S = E \times B/\mu_0$ is the *Poynting vector* that describes the energy flow density of the electromagnetic field. The energy–momentum tensor T obeys a continuity relation: it describes a conserved quantity.

We can sum up by stating that in nature, energy and momentum are conserved, if we take into account the momentum and energy of the electromagnetic field. And the energy–momentum tensor shows again that electrodynamics is a gauge invariant description: the energy and momentum values do not depend on gauge choices.

The energy–momentum tensor, like the Lagrangian, shows that electrodynamics is invariant under *motion inversion*. If all charges change direction of motion – a situation often confusingly called 'time inversion' – they move backwards along the same paths they took when moving forward. Every example of motion due to electric or magnetic

FIGURE 41 Which one is the original landscape? (NOAA).

causes can also take place backwards.

On the other hand, everyday life shows many electric and magnetic effects which are not time invariant, such as the breaking of bodies or the burning of electric light bulbs. Can you explain how this fits together?

We also note that charges and mass destroy a symmetry of the vacuum that we mentioned in special relativity: only the vacuum is invariant under conformal transformations. In particular, only the vacuum is invariant under the spatial inversion $r \rightarrow 1/r$. Any other physical system does not obey conformal symmetry.

To sum up, electrodynamic motion, like all other examples of motion that we have encountered so far, is deterministic, slower than c, reversible and conserved. This is no big surprise. Nevertheless, two other symmetries of electromagnetism deserve special mention.

WHAT IS A MIRROR?

We will study the strange properties of mirrors several times during our walk. We start with the simplest one first. Everybody can observe, by painting each of their hands in a different colour, that a mirror does *not* exchange right and left, as little as it exchanges up and down; however, a mirror does exchange right and left *handedness*. In fact, it does so by exchanging front and back.

Electrodynamics give a second answer: a mirror is a device that switches magnetic north and south poles. Can you confirm this with a diagram?

But is it always possible to distinguish left from right? This seems easy: this text is quite different from a bɘɿoɿɿim version, as are many other objects in our surroundings. But take a simple landscape. Are you able to say which of the two pictures of Figure 41 is the original?

Astonishingly, it is actually impossible to distinguish an original picture of nature from its mirror image if it does not contain any human traces. In other words, everyday nature is somehow left–right symmetric. This observation is so common that all candidate exceptions, from the jaw movement of ruminating cows to the helical growth of plants, such as hops, or the spiral direction of snail shells, have been extensively studied.[*]

---

[*] The most famous is the position of the heart. The mechanisms leading to this disposition are still being investigated. Recent research suggests that the oriented motion of the cilia on embryos, probably in the

Can you name a few more?

The left–right symmetry of nature appears because everyday nature is described by gravitation and, as we will see, by electromagnetism. Both interactions share an important property: substituting all coordinates in their equations by the negative of their values leaves the equations unchanged. This means that for any solution of these equations, i.e., for any naturally occurring system, a mirror image is a possibility that can also occur naturally. Everyday nature thus cannot distinguish between right and left. Indeed, there are right *and* left handers, people with their heart on the left *and* others with their heart on the right side, etc.

To explore further this strange aspect of nature, try the following experiment: imagine you are exchanging radio messages with a Martian; are you able to explain to him what right and left are, so that when you meet, you are sure you are talking about the same thing?

Actually, the mirror symmetry of everyday nature – also called its *parity invariance* – is so pervasive that most animals cannot distinguish left from right in a deeper sense. Most animals react to mirror stimuli with mirror responses. It is hard to teach them different ways to react, and it is possible almost only for mammals. The many experiments performed in this area gave the result that animals have symmetrical nervous systems, and possibly only humans show *lateralization*, i.e., a preferred hand and different uses for the left and the right parts of the brain.

To sum up this digression, classical electrodynamics is left–right symmetric, or parity invariant. Can you show this using its Lagrangian?

Why do metals provide good mirrors? Metals are strong absorbers of light. Any strong absorber has a metallic shine. This is true for metals, if they are thick enough, but also for dye or ink crystals. Any material that strongly absorbs a light wavelength also reflects it efficiently. The cause of the strong absorption of a metal is the electrons inside it; they can move almost freely and thus absorb most visible light frequencies.

Here is a puzzle: a concave mirror shows an inverted image; so does a plane mirror if it is partly folded along the horizontal. What happens if this mirror is rotated around the line of sight?

WHAT IS THE DIFFERENCE BETWEEN ELECTRIC AND MAGNETIC FIELDS?

Obviously, the standard answer is that electric fields have sources, and magnetic fields do not; as a result, magnetic fields are small relativistic effects of importance only when charge velocities are high or when electrical fields cancel out.

For situations involving matter, this clear distinction is correct. Up to the present day, no particle with a magnetic charge, called a *magnetic monopole*, has ever been found, even though its existence is possible in several speculative models of particle physics. If

region called the *node*, determines the right–left asymmetry. The deep origin of this asymmetry is not yet elucidated, however.

Most human bodies have more muscles on the right side for right-handers, such as Albert Einstein and Pablo Picasso, and correspondingly on the left side for left-handers, such as Charlie Chaplin and Peter Ustinov. This asymmetry reflects an asymmetry of the human brain, called lateralization, which is essential to human nature.

Another asymmetry of the human body is the hair whirl on the back of the head; the majority of humans have only one, and in 80 % of the cases it is left turning. But many people have more than one.

found, the action (49) would have to be modified by the addition of a fourth term, namely the magnetic current density. However, no such particle has yet been detected, despite intensive search efforts.

In empty space, when matter is not around, it is possible to take a completely different view. In empty space the electric and the magnetic fields can be seen as two faces of the same quantity, since a transformation such as

$$\begin{aligned} E &\to c\,B \\ B &\to -E/c \end{aligned} \tag{56}$$

called (electromagnetic) *duality* transformation, transforms each vacuum Maxwell equation into the other. The minus sign is necessary for this. (In fact, there are even more such transformations; can you spot them?) Alternatively, the duality transformation transforms F into ${}^*F$. In other words, in empty space we *cannot* distinguish electric from magnetic fields. In particular, it is impossible to say, given a field line in vacuum, whether it is a magnetic or an electric field line.

Matter would be symmetric under duality only if magnetic charges, also called magnetic monopoles, could exist. In that case the transformation (56) could be extended to

$$c\rho_e \to \rho_m \quad , \quad \rho_m \to -c\rho_e \ . \tag{57}$$

For a long time, it was thought that duality can be used in the search for the final, unified theory of physics. However, this hope has evaporated. The reason for this failure can be traced back to a small but ugly fact: the electromagnetic duality transformation changes the sign of the Lagrangian, and thus of the action. Therefore, electromagnetic duality is not a real symmetry of nature, and thus does not help to reach a deeper understanding of electromagnetism.

Duality, by the way, is a symmetry that works *only* in Minkowski space-time, i.e., in space-times of 3 + 1 dimensions. Mathematically, duality is closely related to the existence of quaternions, to the possibility of interpreting Lorentz boosts as rotations in 3 + 1 dimensions, and last, but not least, to the possibility of defining other smooth mathematical structures than the standard one on the space R^4. These mathematical connections are mysterious for the time being; they somehow point to the special role that *four* space-time dimensions play in nature. More details will become apparent in the last volume of our mountain ascent.

Could electrodynamics be different?

We saw that electrodynamics is based on three ideas: the conservation of charge, the speed limit for charges, and Coulomb's relation. Could any of these be wrong or need modification?

Experiments imply that the only candidate for modification is Coulomb's relation. Any interaction, such as Coulomb's relation (4), which acts, for one given observer, between two particles independently of 3-velocity, must depend on 3-velocity for other inertial observers.* Such an interaction must also depend on the 4-velocity, to ensure the

---

* This can be deduced from special relativity, from the reasoning of page 46 or from the formula in the

requirement from special relativity that the 4-acceleration must be 4-orthogonal to the 4-velocity. The simplest case is an interaction in which the acceleration is proportional to the 4-velocity. Together with the request that the interaction leaves the rest mass constant, we then recover electrodynamics. In fact, the requirements of gauge symmetry and of relativistic invariance also make it impossible to modify electrodynamics. In short, it does not seem possible to have a behaviour different from $1/r^2$ for a classical interaction.

Maybe a tiny deviation from Coulomb's relation is possible? An inverse square dependence implies a vanishing mass of light and light particles, the photons. Is the mass really zero? The issue has been extensively studied. A massive photon would lead to a wavelength dependence of the speed of light in vacuum, to deviations from the inverse square 'law', to deviations from Ampère's 'law', to the existence of longitudinal electromagnetic waves and more. No evidence for these effects has ever been found. A summary of these studies shows that the photon mass is below 10^{-53} kg, or maybe 10^{-63} kg. Some arguments are not universally accepted, thus the limit varies somewhat from researcher to researcher.

A small non-vanishing mass for the photon would change electrodynamics somewhat. The inclusion of a tiny mass poses no special problems, and the corresponding Lagrangian, the so-called *Proca Lagrangian*, has already been studied, just in case.

Strictly speaking, the photon mass cannot be said to vanish. In particular, a photon with a Compton wavelength of the radius of the visible universe cannot be distinguished from one with zero mass through any experiment. This gives a limit mass of 10^{-69} kg for the photon. Photons with such a small mass value would not invalidate electrodynamics as we know it. We note that the experimental limits are still much larger.

Interestingly, a non-zero mass of the photon would imply the lack of magnetic monopoles, as the symmetry between electric and magnetic fields would be broken. It is therefore important on the one hand to try to improve the experimental mass limit for photons, and on the other hand to explore whether the limit due to the universe's size has any implications for this issue. The question is still open.

In summary, it seems extremely difficult to find modifications of elctrodynamics that do not disagree with experiment. Electrodynamics is fixed once for all.

The brain: the toughest challenge for electrodynamics

Researchers working on classical electrodynamics still face a fascinating experimental and theoretical issue: understanding the process of thought. Researchers face two challenges in this domain. First, they must find ways to *model* the thought process. Second, the technology to *measure* the currents in the brain must be extended. In both domains, recent progress has been spectacular.

Important research has been carried out on many levels of thought modelling. For example, research using computer tomography, PET scans and MRI imaging has shown that the distinction between the *conscious* and the *unconscious* can be measured: it has a biological basis. Psychological concepts, such as *repression* of unpleasant thoughts, can actually be observed in brain scans. Modellers of brain mechanisms are learning to take some of the concepts of psychology as descriptions for actual physical processes. This research approach is still in its infancy, but very promising.

footnote of page 74.

Ref. 43

Ref. 44

Ref. 44
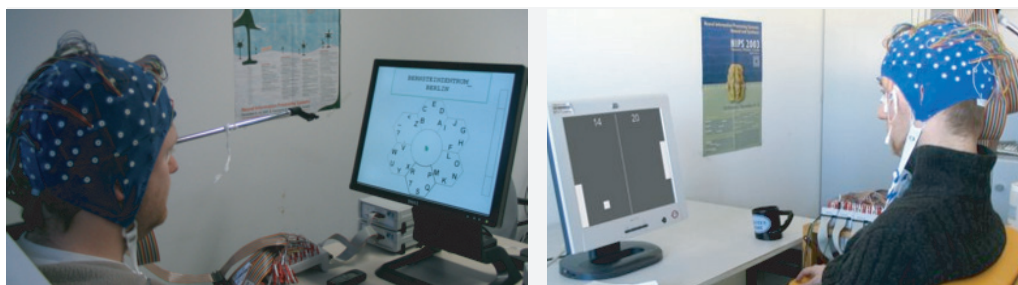
FIGURE 42 Typing a letter and playing video tennis using thought alone (© Fraunhofer FIRST).

About the specific aspects of the working of the brain, such as learning, storage, recognition of shapes, location of sound sources or map formation, modern neurobiology and animal experimentation have allowed to deduce models that allow quantitative predictions. More on this will be told below.

On the experimental side, research into magnetoencephalography devices is making rapid progress. The magnetic fields produced by brain currents are as low as 10 fT, which require sensors at liquid helium temperature and a good shielding of background noise. Improving the sensitivity and the spatial resolution of these systems is a central task. Also computer models and algorithms are making rapid progress.

The whole programme would be complete as soon as, in a distant future, a sensitive measuring apparatus could detect what is going on inside the brain and then could deduce or 'read' the thoughts of a person from these measurements. Thought reading might be the most complex of all challenges that science is facing. Clearly, such a feat will require involved and expensive machinery, so that there is no danger for a misuse of the technique. (There are also good reasons to believe that actual thought reading will never be possible in this way, due to the lack of localization of cognitive thought inside the brain and due to the variations in cognitive processing from one person to another.) But the understanding and modelling of the brain will be a useful technology in numerous aspects of daily life, especially for the disabled.

On the path towards thought reading, the small progress that has been achieved so far is already fascinating. Wearing a cap full of electric contacts (thus without any surgery on the brain) and looking at a computer screen, it is now possible to type letters using the power of thought alone. Such a system is shown in Figure 42. The user controls the computer simply by *imagining* that he turns the arrow on the screen with his right hand. The brain currents created by the imagination process are read out and translated into computer commands by an electronic device. The system, based on neural network algorithms, works after only 20 minutes of training with a particular person. In this way, the system allows people who are fully paralysed to communicate with others again. The system is so fast that it allows playing 'mental video tennis' on a computer screen.

Typing with thought alone is possible because the brain region responsible for the hand is near the skull, so that signals for hand rotation can be read out with sufficient spatial resolution by the electrodes on the cap. Researchers know that resolution limitations do not allow reading out the commands for single fingers in this way. For such high resolution tasks, electrodes still need to be *implanted* inside the relevant brain region. However, at present the functional lifetime for such electrodes is only a few months, so

that the dream of controlling machines or even artificial limbs in this way is still distant.

CHALLENGES AND FUN CURIOSITIES ABOUT ELECTRODYNAMICS

Not only animals, also plants can feel electric and magnetic fields. At least for magnetic fields, the sensors seem to use very similar mechanisms to those used by animals and bacteria.

* *

If you calculate the Poynting vector for a charged magnet – or simpler, a point charge near a magnet – you get a surprising result: the electromagnetic energy flows in circles around the magnet. How is this possible? Where does this angular momentum come from?

Challenge 90 s

Worse, any atom is an example of such a system – actually of two such systems. Why is this effect not taken into account in calculations in quantum theory?

Ref. 46

* *

Perfectly spherical electromagnetic waves are impossible in nature. Can you show this using Maxwell's equation of electromagnetism, or even without them?

Challenge 91 s

* *

Can you sketch the Poynting field along a double wire? Around a double wire with resistance?

Challenge 92 ny

* *

Can you sketch the Poynting field in a transformer?

Challenge 93 ny

* *

There is a famous ambiguity in the definition of electromagnetic energy. Richard Feynman talks about it in his *Lectures on Physics*, in section 27-4. Can you find an experiment where this ambiguity plays a role?

Challenge 94 ny

* *

Any wall plug is a dipole driven by an alternating electric field. Why does a wall plug, delivering 230 V or 100 V at 50 Hz or 60 Hz, not radiate electromagnetic fields?
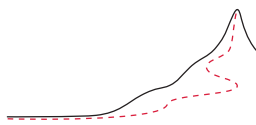
Challenge 95 s

SUMMARY

In summary, the electromagnetic field carries energy, linear momentum and angular momentum. It is thus appropriate to say that the electromagnetic field *moves*. The motion of the electromagnetic field can be visualized as the motion of its electric and its magnetic field lines. The motion of the electromagnetic field is described by a least action principle.

We are directly lead to ask: what then is the nature of light?

WHAT IS LIGHT?
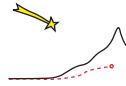
T HE nature of light has fascinated explorers of nature since at least the time of
the ancient Greeks. The answer appeared in 1848, when Gustav Kirchhoff noted
that the experimental values on both sides of the equation

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \ .\tag{58}$$

agreed within measurement errors. This suggested the answer to the question two thou-
sand years earlier: *light is an electromagnetic wave*. Ten years later, in 1858, Bernhard
Riemann* proved mathematically that any electromagnetic wave must propagate with
a speed c given by the above equation. Note that the right-hand side contains electric
and magnetic quantities, and the left-hand side is an optical quantity. The expression of
Kirchhoff and Riemann thus unifies electromagnetism and optics. The modern value for
the speed of electromagnetic waves, usually called c from Latin *celeritas*, is

$$c = 299\,792\,458\,\text{m/s} \ .\tag{59}$$

The value for c is an integer number, because the meter is nowadays defined in such a
way as to exactly achieve this number.

In 1865, Maxwell summarized all data on electricity and magnetism collected in the
2500 years in his equations. Almost nobody read his papers, because he wrote them us-
ing quaternions. The equations were then simplified independently by Heinrich Hertz
and Oliver Heaviside. They deduced the original result of Riemann: in the case of empty
space, the equations of the electromagnetic potentials can be written as

$$\Box\,\boldsymbol{A} = 0 \quad \text{or, equivalently} \quad \varepsilon_0\mu_0\frac{\partial^2\varphi}{\partial t^2} + \frac{\partial^2 A_x}{\partial x^2} + \frac{\partial^2 A_y}{\partial y^2} + \frac{\partial^2 A_z}{\partial z^2} = 0 \ .\tag{60}$$

This evolution equation is a *wave equation*, because it admits solutions of the type

$$\boldsymbol{A}(t,\boldsymbol{x}) = \boldsymbol{A}_0\sin(\omega t - \boldsymbol{k}\boldsymbol{x} + \delta) = \boldsymbol{A}_0\sin(2\pi f t - 2\pi\boldsymbol{x}/\lambda + \delta) \ ,\tag{61}$$

---

* Bernhard Riemann (b. 1826 Breselenz, d. 1866 Selasca), important German mathematician. A genial math-
ematician, he also studied curved space, providing several of the mathematical and conceptual foundations
of general relativity, but then died at an early age.
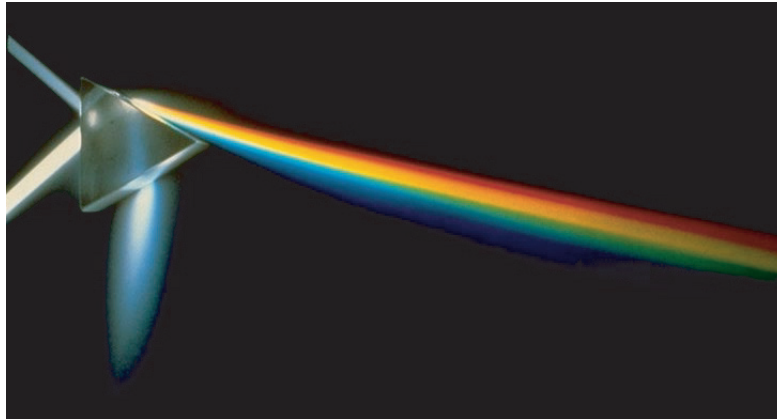
which are commonly called harmonic *plane electromagnetic waves*. We recall that a *wave* in physics is any propagating imbalance, and that a *harmonic* wave is a wave described by a sine curve.

Such a harmonic plane electromagnetic wave satisfies equation (60) for any value of *amplitude A_0*, of *phase δ*, and of *angular frequency ω*, provided the angular frequency and the *wave vector $\boldsymbol{k}$* satisfy the relation

$$\omega(\boldsymbol{k}) = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} k \quad \text{or} \quad \omega(\boldsymbol{k}) = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \sqrt{\boldsymbol{k}^2} \ . \tag{62}$$

The relation $\omega(\boldsymbol{k})$ between the angular frequency and the wave vector, the so-called *dispersion relation*, is the main property of any type of wave, be it a sound wave, a water wave, an electromagnetic wave, or any other kind.

The specific dispersion relation (62) is *linear* and implies a *phase velocity*, the velocity with which wave crests and troughs move, given by $\omega/k = 1/\sqrt{\varepsilon_0 \mu_0} = c$, thus reproducing the result by Kirchhoff and Riemann.

In empty space, experiments confirm that the phase velocity c is independent of the frequency of the wave. The phase velocity thus characterizes electromagnetic waves, and distinguishes them from all other types of waves in nature.

What are electromagnetic waves?

To get a clearer idea of electromagnetic waves, we explore their properties. The wave equation (60) for the electromagnetic field is *linear* in the field; this means that the sum of two situations allowed by it is itself an allowed situation. Mathematically speaking, any *superposition* of two solutions is also a solution. We therefore know that electromagnetic waves must show *interference*, as all linear waves do.

Linearity implies that two waves can cross each other without disturbing each other, and that electromagnetic waves can travel undisturbed across static electromagnetic fields.

Linearity also means that every electromagnetic wave can be described as a super-
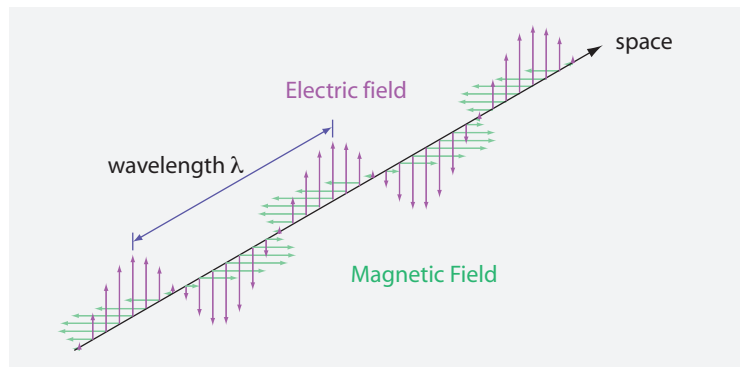
FIGURE 45 A plane, monochromatic and linearly polarized electromagnetic wave, showing the evolution of the electric field, the magnetic field, and again the electric field, in a further visualization (Mpg films © Thomas Weiland).

FIGURE 46 Heinrich Hertz (1857–1894).

position of harmonic, or pure sine waves, each of which is described by expression (61). The simplest possible electromagnetic wave, the harmonic plane wave with linear polarization, is illustrated in Figure 44. Note that for this simplest type of waves, the electric and the magnetic field are *in phase*. (Can you prove this experimentally and by calculation?) The surfaces formed by all points of maximal field intensity are parallel planes, spaced by (half the) wavelength; these planes move along the direction of the propagation with the phase velocity.

After Riemann and Maxwell predicted the existence of electromagnetic waves, in the years between 1885 and 1889, Heinrich Hertz* discovered and studied them. He fabri-

---

* Heinrich Rudolf Hertz (b. 1857 Hamburg, d. 1894 Bonn), important Hamburger theoretical and experimental physicist. The unit of frequency is named after him. Despite his early death, Hertz was a central
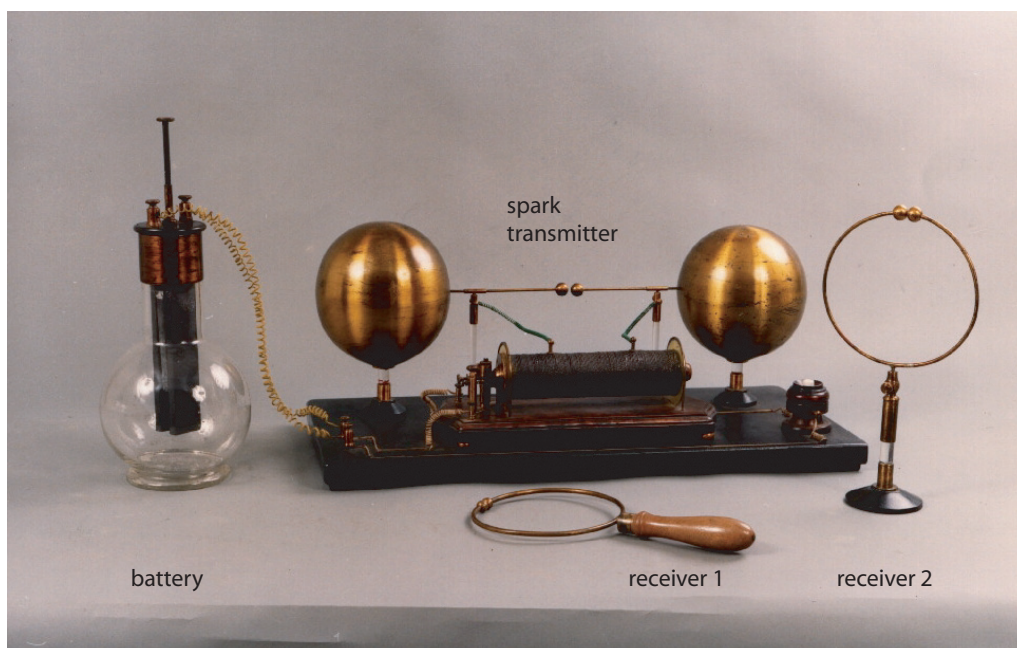
FIGURE 47 A reconstruction of one of the first transmitters and receivers of electromagnetic waves by Heinrich Hertz (© Fondazione Guglielmo Marconi).

cated a very simple transmitter and receiver for 2 GHz waves, shown in Figure 47. Such waves are still used today: cordless telephones and the last generation of mobile phones work at this frequency – though the transmitters and the receivers look somewhat differently nowadays. Such waves are now also called *radio waves*, since physicists tend to call all moving force fields *radiation*, recycling somewhat incorrectly a Greek term that originally meant 'light emission.'

Today Hertz's experiment can be repeated in a much simpler way. As shown in Figure 48, a budget of a few euro is sufficient to remotely switch on a light emitting diode with a gas lighter. (After each activation, the coherer has to be gently tapped, in order to get ready for the next activation.) Attaching longer wires as antennas and ground allows this set-up to achieve transmission distances up to 30 m.

Hertz also measured the *speed* of the waves he produced. In fact, you can also measure the speed at home, with a chocolate bar and a (older) kitchen microwave oven. A microwave oven emits radio waves at 2.5 GHz – not far from Hertz's value. Inside the oven, these waves form standing waves. Just put the chocolate bar (or a piece of cheese) in the oven and switch the power off as soon as melting begins. You will notice that the bar melts at regularly spaced spots. These spots are half a wavelength apart. From the measured wavelength value and the frequency, the speed of light and radio waves simply follows as the product of the two.

If you are not convinced, you can measure the speed directly, by telephoning a friend on another continent, if you can make sure of using a satellite line (choose a low cost

figure in the development of electromagnetism, in the explanation of Maxwell's theory and in the unfolding of radio communication technology. More about him on .
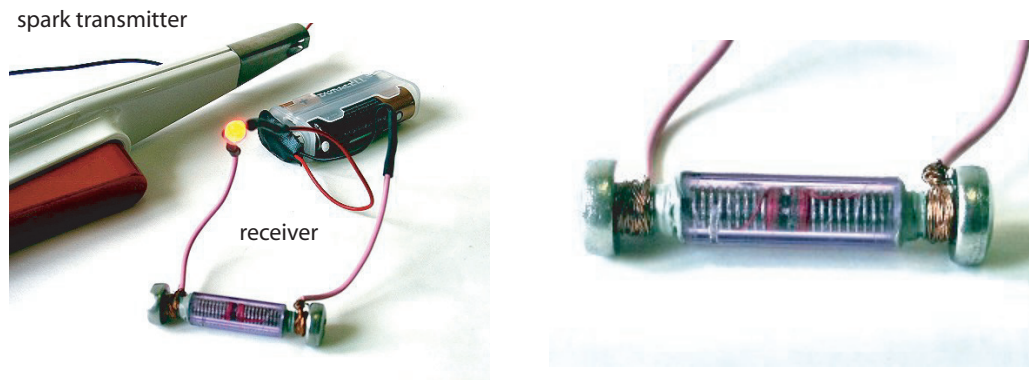
FIGURE 48 The simplest radio transmitter possible, a gas lighter and a wire, together with the simplest radio receiver possible, built from a battery pack, a light emitting diode, and a simple *coherer* made from a ball pen housing, two screws and some metal powder (© Guido Pegna).

provider). There is about half a second additional delay between the end of a sentence and the answer of the friend, compared with normal conversation. In this half second, the signal goes up to the geostationary satellite, down again and returns the same way. This half second gives a speed of $c \approx 4 \cdot 36\,000\,\text{km}/0.5\,\text{s} \approx 3 \cdot 10^5\,\text{km/s}$, which is close to the precise value. Radio amateurs who reflect their signals from the Moon can perform the same experiment and achieve higher precision.

In summary, electromagnetic waves exist and move with the speed of light.

Light as an electromagnetic wave

But the electromagnetic wave equation is much more interesting. The wave equation confirmed earlier predictions that *light* itself is an electromagnetic wave, albeit with a much higher frequency and much shorter wavelength. Let us see how we can check this.

It is easy to confirm the *wave* properties of light; indeed they were known already long before Maxwell. In fact, the first to suggest that light is a (kind of) wave was, around the year 1678, the important physicist Christiaan Huygens.* You can confirm that light is a wave with your own fingers. Simply place your hand one or two centimetres in front of your eye, look towards the sky through the gap between the middle and the index finger and let the two fingers almost touch. You will see a number of dark lines crossing the gap. These lines are the interference pattern formed by the light behind the slit created by the fingers. *Interference* is the name given to the amplitude patterns that appear when several waves superpose.** The interference patterns depend on the spacing between the fingers. This experiment therefore allows you to estimate the wavelength of light, and thus, if you know its speed, its frequency. Can you do this?

Challenge 98 s

---

* Christiaan Huygens (b. 1629 's Gravenhage, d. 1695 Hofwyck) was one of the main physicists and mathematicians of his time. Huygens clarified the concepts of mechanics; he also was one of the first to show that light is a wave. He wrote influential books on probability theory, clock mechanisms, optics and astronomy. Among other achievements, Huygens showed that the Orion Nebula consists of stars, discovered Titan, the moon of Saturn, and showed that the rings of Saturn consist of rock. (This is in contrast to Saturn itself, whose density is lower than that of water.)

Challenge 97 s ** Where does the energy go in an interference pattern?

FIGURE 49 The primary and secondary rainbow, and the supernumerary bows below the primary bow (© Antonio Martos and Wolfgang Hinz).

Historically, another effect was central in convincing everybody that light was a wave: supernumerary rainbows, the additional bows below the main or primary rainbow. If we look carefully at a rainbow, below the main red–yellow–green–blue–violet bow, we observe weaker, additional green, blue and violet bows. Depending on the intensity of the rainbow, several of these supernumerary rainbows can be observed. They are due to interference of light triggered by the water droplets, as Thomas Young showed around 1803.* Indeed, the repetition distance of the supernumerary bows depends on the radius of the average water droplets that form them. (Details about the normal rainbows are given below.) Supernumerary rainbows were central in convincing people that light is a wave. It seems that in those times scientists either did not trust their own fingers, or did not have any.

There are many other ways in which the wave character of light can be made apparent. Maybe the most beautiful is an experiment carried out by a team of Dutch physicists in 1990. They simply measured the light transmitted through a *slit* in a metal plate. It turns out that the transmitted intensity depends on the width of the slit. Their surprising result is shown in Figure 50. Can you explain the origin of the unexpected intensity steps in the curve?

Numerous other experiments on the creation, detection and measurement of electromagnetic waves were performed between the seventeenth and the twentieth century. For example, in 1800, William Herschel discovered *infrared light* using a prism and a thermometer. (Can you guess how?) In 1801, Johann Wilhelm Ritter (1776–1810) a more than colourful figure of natural Romanticism, discovered *ultraviolet light* using silver chloride, AgCl, and again a prism.

---

* Thomas Young (1773 Milverton–1829), read the bible at two, spoke Latin at four; a doctor of medicine, he became a professor of physics. He introduced the concept of *interference* into optics, explaining Newtonian rings and supernumerary rainbows; he was the first person to determine light's *wavelength*, a concept that he also introduced, and its dependence on colour. He was the first to deduce the three-colour vision explanation of the eye and, after reading of the discovery of polarization, explained light as a transverse wave. In short, Young discovered most of what people learn at secondary school about light. He was a universal talent: he also worked on the deciphering of hieroglyphs, studied languages and introduced the term 'Indo-European', explored ship building and many engineering problems. Young collaborated with Fraunhofer and Fresnel. In Britain his ideas on light were not accepted, since Newton's followers crushed all opposing views. Towards the end of his life, his results were finally made known to the physics community by Fresnel and Helmholtz.
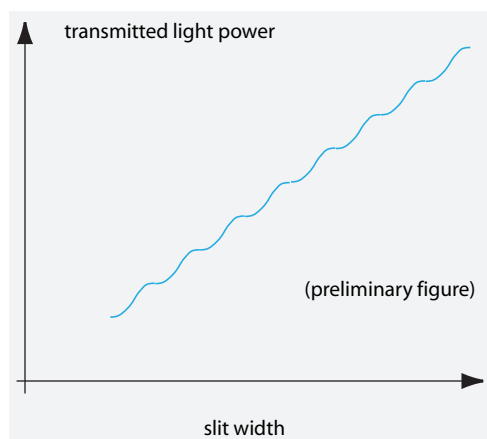
FIGURE 50 The light power transmitted through a slit as function of its width.

FIGURE 51 Two Gaussian beams interfering at an angle (© Rüdiger Paschotta).

The result of all these experiments is that electromagnetic waves, including light, can be primarily distinguished by their wavelength or frequency. The main categories are listed in Table 13. For visible light, the wavelength lies between 0.4 μm (violet) and 0.8 μm (red). The wavelength of light determines its *colour*.

At the end of the twentieth century the final confirmation of the wave character of light became possible. Using quite sophisticated experiments. researchers measured the oscillation frequency of light *directly*. The value, between 375 and 750 THz, is as predicted. The value is so high that its detection was impossible for a long time. But with these modern experiments the dispersion relation (62) of light has finally been confirmed in all its details.

Polarization and electromagnetic waves

We are left with one additional question about light. If light oscillates, in which direction does this occur? The answer is hidden in the parameter A_0 in expression (61), but shown in Figure 44 and Figure 45. The fields in electromagnetic waves oscillate in directions *perpendicular* to their motion. Therefore, even for identical frequency and phase, waves can still differ: they can have different *polarization* directions. For example, the polarization of radio transmitters determines whether radio antennas of receivers have to be kept horizontal or vertical. Also for light, polarization is easily achieved, e.g. by shining it

primary
infrared
rainbow

secondary
infrared
rainbow

secondary
visible
rainbow

primary
visible
rainbow

FIGURE 52 The same rainbow
in the visible and in the infrared,
showing how infrared comes
before red (© Stefan Zeiger).

through a stretched plastic film. When the polarization of light was discovered in 1808 by
the French physicist Louis Malus (1775–1812), it definitively established the wave nature
of light. Malus discovered it when he looked at the strange double images produced by
calcite, a transparent crystal found in many minerals. Calcite ($CaCO_3$) splits light beams
into two – it is *birefringent* – and polarizes them differently. That is the reason that calcite
is part of every crystal collection.

    If you ever get hold of a piece of transparent calcite, do look through it at something
written on paper. Figure 53 shows two examples. (Can you show that *trirefringence*, if
defined as the appearance of three images, cannot exist?)

    By the way, the human eye is almost unable to detect polarization, in contrast to the
eyes of many insects, spiders and certain birds. Honey bees use polarization to deduce
the position of the Sun, even when it is hidden behind clouds, and use the effect for
navigation. Some beetles of the genus *Scarabeus* even use the polarization of moonlight
for navigation, and many insects use polarization of sunlight to distinguish water sur-
faces from mirages. (Can you find out how?) But in 1844, the Austrian mineralogist Wil-
helm Haidinger discovered that the human eye has the same ability: there is a way to
observe the polarization of light with the unaided human eye. The best way to observe
the effect is by looking at a distance of about an arms's length on a white LCD screen
and slowly tilt your head. You will note an *extremely faint* yellow or yellow-blue pattern,
about two finger's wide, that is superimposed on the white background. This pattern is
called *polarization brush* or *Haidinger's brush*. A rough illustration is given in Figure 54.
The weak effect disappears after a few seconds if the head stops rotating along the line of
sight. Haidinger's brush is due to the birefringence of the cornea and the lens of the hu-
man eye, together with the morphology of the macula lutea. The cornea acts as a radially
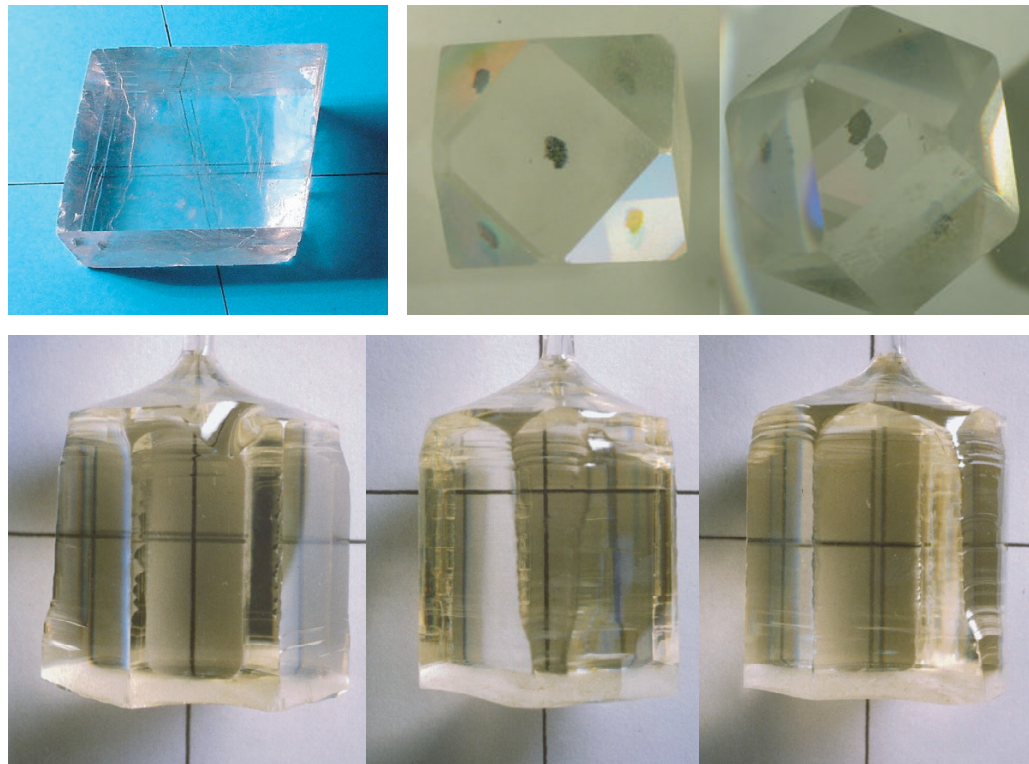
FIGURE 53 Birefringence in crystals: *calcite* lying on crossed lines (top left, crystal size around 4 cm), *rutile* lying on an ink spot, photographed along the optical axis (middle) and at an angle to it (top right, crystal size around 1 cm), and an octagonal sodium vanadate crystal doped with manganese, showing *three* different behaviours (bottom, crystal diameter 1.9 cm) (© Roger Weller/Cochise College, Brad Amos, Martin Pietralla).
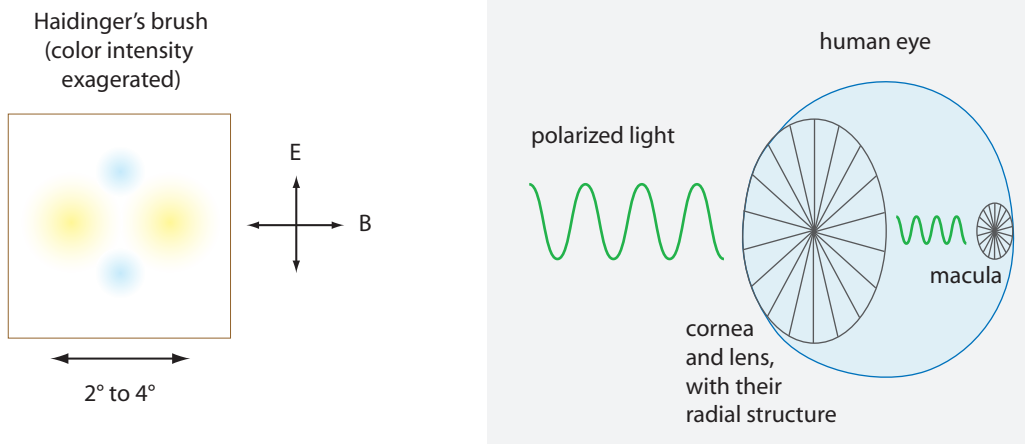
FIGURE 54 Haidinger's brush and its origin in the human eye.

oriented, colour-dependent polarizer, whereas the yellow spot acts as a radially oriented analyser. In short, the human eye is indeed able to see the directions in which the electric
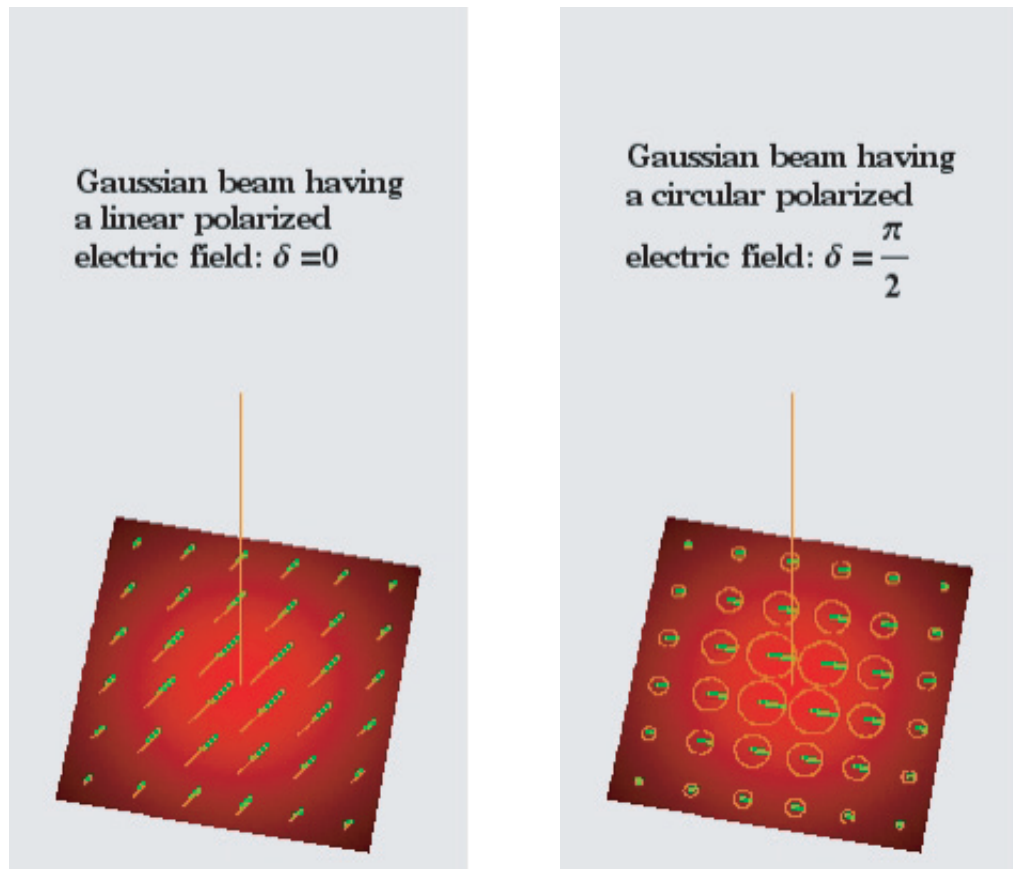
FIGURE 55 Left: the electric field of a Gaussian, linearly polarized electromagnetic wave (a beam); right: a Gaussian, circularly polarized beam (Quicktime film © José Antonio Díaz Navas).

and magnetic field of light are oscillating.

Haidinger's brush, being yellow, is also visible in the blue sky, provided that the air is clear. (Indeed, it is easily drowned out by multiple scattering, and therefore provides a test of atmospheric transparency.) In the sky, Haidinger's brush is barely the size of a thumbnail at arm's length. (The angular size is the angular size of the macula.) The yellow arm of the cross points to the Sun, if you look about 90° away from it, high in the sky. To see it really clearly, hold a polaroid (or polaroid sunglasses) up to look through, and rotate it about the line of sight.

Note that all possible polarizations of light form a continuous set. However, a general plane wave can be seen as the superposition of two orthogonal, linearly polarized waves with different amplitudes and different phases. Most books show pictures of plane, linearized electrodynamic waves. Essentially, electric fields look like water waves generalized to three dimensions, the same for magnetic fields, and the two are perpendicular to each other. Can you confirm this?

Interestingly, a generally polarized plane wave can also be seen as the superposition of right and left *circularly polarized waves*. An illustration of a circularly polarized wave

is given in Figure 55.

LIGHT AND OTHER ELECTROMAGNETIC WAVES

The experiments so far show that electromagnetic waves exist and move with the same speed as light. To confirm that light waves are indeed *electromagnetic* is more difficult. The most convincing proof would be to repeat Hertz's experiments for light. In Hertz's experiment, shown in Figure 47, the receiver is a simple open metal circle; when the wave – more precisely, its magnetic field – arrives, a spark is generated and the wave is thus detected. In an almost incredible feat of miniaturization, in 2009, the research group of Kobus Kuipers managed to make metal rings much smaller than a micrometre, and repeat the experiment for light. They could clearly discern the maxima and minima of waves, as well as their polarization. They thus showed that light is an electromagnetic wave in exactly the same way as Hertz did for radio waves.

Ref. 55

Of course, people in the 19th century had less technology at their disposal and were not easily convinced. They had to look for other ways to show that light is electromagnetic in nature. Now, since the evolution equations of the electrodynamic field are linear, additional electric or magnetic fields alone do not influence the motion of light. On the other hand, we know that electromagnetic waves are emitted only by accelerated charges, and that all light is emitted from matter. It thus follows that matter is full of electromagnetic fields and accelerated electric charges. This in turn implies that the influence of matter on light can be understood from its internal electromagnetic fields and, in particular, that subjecting matter to an *external* electromagnetic field should change the light it emits, the way matter interacts with light, or generally, the material properties as a whole.

Challenge 104 e

Searching for effects of electricity and magnetism on matter has been a main effort of physicists for over a hundred years. For example, electric fields influence the light transmission of oil, an effect discovered by John Kerr in 1875.* Also the discovery that certain gases change colour when subject to a field yielded several Nobel Prizes for physics. With time, many more influences on light-related properties by matter subjected to fields were found. An extensive list is given in the table on page 158. It turns out that apart from a few exceptions the effects can *all* be described by the electromagnetic Lagrangian (49), or equivalently, by Maxwell's equations (53). In summary, classical electrodynamics indeed unifies the description of electricity, magnetism and optics; all phenomena in these fields, from the rainbow to radio and from lightning to electric motors, are found to be different aspects of the evolution of the electromagnetic field.

Page 74

After two centuries of research, it became clear that light is only a very small section of the full *spectrum of electromagnetic waves*, which contains the waves from the smallest possible to the largest possible wavelengths. The full spectrum is given in the following table.

TABLE 13 The electromagnetic spectrum.

FREQUENCY	WAVELENGTH	NAME	MAIN PROPERTIES	APPEARANCE	USE
$3 \cdot 10^{-18}$ Hz	10^{26} m	**Lower frequency limit**		see the section on cosmology	

* John Kerr (1824–1907), Scottish physicist, friend and collaborator of William Thomson.

FRE-QUENCY	WAVE-LENGTH	NAME	MAIN PROPERTIES	APPEARANCE	USE
< 10 Hz	> 30 Mm	**Quasistatic fields**		intergalactic, galactic, stellar and planetary fields, brain, electrical fish	power transmission, accelerating and deflecting cosmic radiation
		Radio waves		electronic devices	
10 Hz– 50 kHz	30 Mm– 6 km	ELW	go round the globe, penetrate into water, penetrate metal	nerve cells, electromechanical devices	power transmission, communication through metal walls, communication with submarines www.vlf.it
50 – 500 kHz	6 km– 0.6 km	LW	follow Earth's curvature, felt by nerves ('bad weather nerves')	emitted by thunderstorms	radio communications, telegraphy, inductive heating
500 – 1500 kHz	600 m– 200 m	MW	reflected by night sky		radio
1.5 – 30 MHz	200 m–10 m	SW	circle world if reflected by the ionosphere, destroy hot air balloons	emitted by stars	radio transmissions, radio amateurs, spying
15 – 150 MHz	20 m–2 m	VHF	allow battery operated transmitters	emitted by Jupiter	remote controls, closed networks, tv, radio amateurs, radio navigation, military, police, taxi
150 – 1500 MHz	2 m–0.2 m	UHF	*idem*, line of sight propagation		radio, walkie-talkies, tv, mobile phones, internet via cable, satellite communication, bicycle speedometers
		Microwaves			
1.5 – 15 GHz	20 cm–2 cm	SHF	*idem*, absorbed by water	night sky, emitted by hydrogen atoms	radio astronomy, used for cooking (2.45 GHz), telecommunications, radar
15 – 150 GHz	20 mm– 2 mm	EHF	*idem*, absorbed by water		
		Infrared	allows night vision	emitted by every warm object	satellite photography of Earth, astronomy

FRE-QUENCY	WAVE-LENGTH	NAME	MAIN PROPERTIES	APPEARANCE	USE
0.3 – 100 THz	1000 –3 μm	IRC or far infrared		sunlight, living beings	seeing through clothes, envelopes and teeth
100 – 210 THz	3 μm– 1.4 μm	IRB or medium infrared		sunlight	used for optical fibre communications for telephone and cable TV
210 – 384 THz	1400– 780 nm	IRA or near infrared	penetrates for several cm into human skin	sunlight, radiation from hot bodies	healing of wounds, rheumatism, sport physiotherapy, hidden illumination
375 – 750 THz	800– 400 nm	**Light**	not (much) absorbed by air, detected by the eye (up to over 900 nm at sufficient power)	heat ('hot light'), lasers & chemical reactions e.g. phosphor oxidation, fireflies ('cold light')	definition of straightness, enhancing photosynthesis in agriculture, photodynamic therapy, hyperbilirubinaemia treatment
384 – 484 THz	780–620 nm	Red	penetrate flesh	blood	alarm signal, used for breast imaging Ref. 56
	700 nm	Laboratory primary red		filtered tungsten lamp	colour reference for printing, painting, illumination and displays
484 – 511 THz	620–587 nm	Orange		various fruit	attracts birds and insects
511 – 525 THz	587–571 nm	Yellow		majority of flowers	*idem*; best background for reading black text
525 – 614 THz	571–488 nm	Green	maximum eye sensitivity	algae and plants	highest luminous efficiency response ('felt brightness') per light energy for the human eye
	546.1 nm	Laboratory primary green		mercury lamp	colour reference
614 – 692 THz	488–433 nm	Blue		sky, gems, water	
	435.8 nm	Laboratory primary blue		mercury lamp	colour reference
692 – 789 THz	433–380 nm	Indigo, violet		flowers, gems	
		Ultraviolet			

Fre-quency	Wave-length	Name	Main properties	Appearance	Use
789 – 952 THz	380–315 nm	UVA	penetrate 1 mm into skin, darken it, produce vitamin D, suppress immune system, cause skin cancer, destroy eye lens	emitted by Sun, stars and flames	seen by certain birds, integrated circuit fabrication
0.95 – 1.07 PHz	315–280 nm	UVB	*idem*, destroy DNA, cause skin cancer	*idem*	*idem*
1.07 – 3.0 PHz	280–100 nm	UVC	form oxygen radicals from air, kill bacteria, penetrate 10 μm into skin	emitted by Sun, stars and welding arcs	disinfection, water purification, waste disposal, integrated circuit fabrication
3 –24 PHz	100–13 nm	EUV			sky maps, silicon lithography
		X-rays	penetrate materials	emitted by stars, plasmas and black holes	imaging human tissue
24 – 240 PHz	13–1.3 nm	Soft X-rays	*idem*	synchrotron radiation	*idem*
> 240 PHz or > 1 keV	< 1.2 nm	Hard X-rays	*idem*	emitted when fast electrons hit matter	crystallography, structure determination
> 12 EHz or > 50 keV	< 24 pm	γ-rays	*idem*	radioactivity, cosmic rays	chemical analysis, disinfection, astronomy
$2 \cdot 10^{43}$ Hz	$\approx 10^{-35}$ m	**Planck limit**		see last volume of this series	

The slowness of progress in physics

Gustav Kirchhoff's and Bernhard Riemann's expression from the 1850s for the speed of light and all other electromagnetic waves

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \qquad (63)$$

is so strange that we should be intrigued whenever we see it. Something essential is missing. The expression states that the speed c is *independent* of the proper motion of the observer measuring the electromagnetic field and *independent* of the speed of the emit-

ting source. In other words, the speed of light is predicted to be independent of the lamp speed and independent of the observer speed. This is indeed confirmed by all experiments.

In addition, no observer can outrun light. In other words, light does *not* behave like a stream of bullets: the speed of bullet depends on the speed of the gun and of the target. A target can always outrun a bullet, if it moves rapidly enough. The speed of light is a *limit speed*.

Experiments confirm that also the speed of radio waves, of X-rays or of γ-rays is independent of the transmitter and the receiver and has the same value as the speed of light. All this is contained in expression (63).

In short, the speed c is *invariant* and is the *limit energy speed* in nature. Incredibly, *nobody* explored the consequences of this invariance until Lorentz and a few others started doing so in the 1890s. The theory of relativity remained undiscovered for two generations! As in so many other cases, the progress of physics was much slower than necessary.

The invariance of the speed of light is the essential point that distinguishes special relativity from Galilean physics. Since every electromagnetic device – such as every electric motor – makes use of expression (63), every electromagnetic device is a working proof of special relativity.

ANOTHER LOOK AT ELECTROMAGNETIC RADIATION

Electromagnetic waves of lower frequency are commonly used to transmit mobile phone signals, and television, radio and satellite programs. Like light, radio waves are due to moving electrons. In everyday life, light is (usually) generated by electrons accelerated inside atoms or molecules. Radio waves, which have lower frequency and thus larger wavelength, are more easily generated by electrons that are accelerated in metals roughly of the size of the wavelength; such pieces of metal are called *antennas*.

Radio waves emitted by a hand-held device can carry signals round the Earth. In other words, radio waves have a large range. How is this possible? It turns out that the field strength of radio waves decreases as $1/r$, where r is the distance from the source. The field strength thus decreases much more slowly than for static fields, which decrease as Ref. 57 $1/r^2$. Why is this the case?

The slow $1/r$ dependence of radio waves can be understood qualitatively from the drawing shown in Figure 56. It shows the electric field around a charged particle that undergoes the simplest possible accelerated motion: a bounce on a wall. In fact, the last, lower diagram is sufficient to show that the transverse field, given by the kink in the Challenge 105 d   electric field lines, decreases as $1/r$. Can you deduce the dependence?

If we perform the construction of the field lines for a charge that undergoes *repeated* bounces, we get field lines with regularly spaced kinks that move away from the source. For a charge undergoing *harmonic* motion, we get the field lines shown in Figure 57. The figure thus shows the mechanism of the simplest antenna (or light source) one can imagine.

The magnitude of the transverse electric field can also be used to deduce the relation between the acceleration a of a charge q and the radiated electromagnetic power P. First, the transverse electric field (calculated in the last challenge) has to be squared, to give the local electric energy density. Then it has to be doubled, to include magnetic energy.

circle radius is ct, where t is the time
since the bounce took place

actual charge
position

wall

charge position
had it not
bounced

electrical field lines

complete
electrical field lines
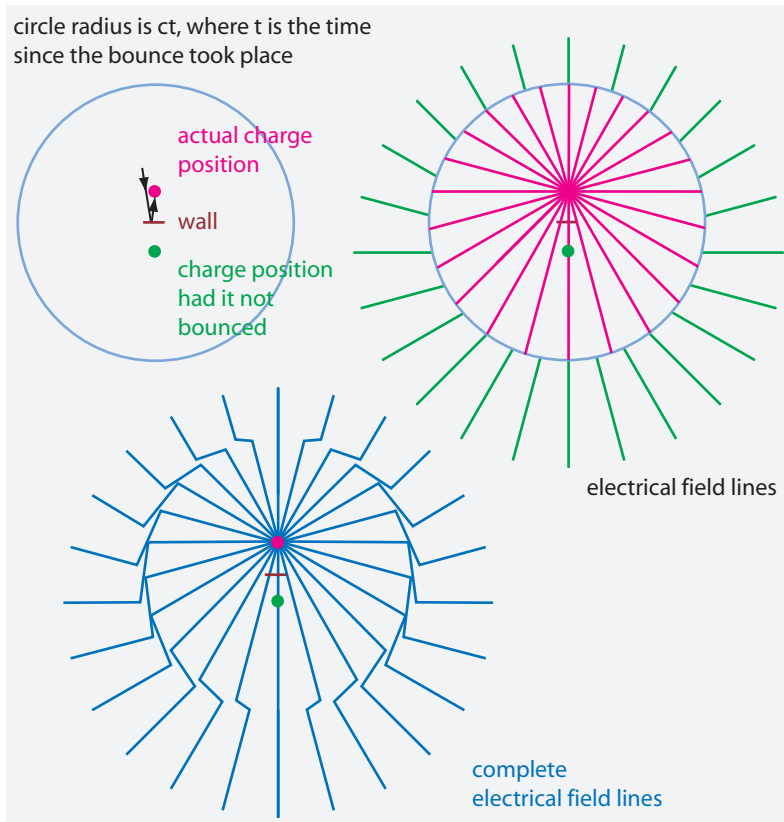
FIGURE 56
Constructing, in three
steps, the electrical
field around a
charged particle
bouncing from a wall.
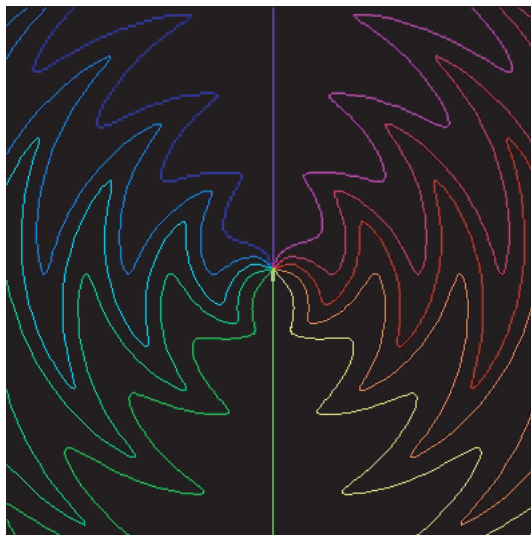
FIGURE 57 The electrical field around a
particle oscillating in vertical direction
(QuickTime film © Daniel Schroeder).
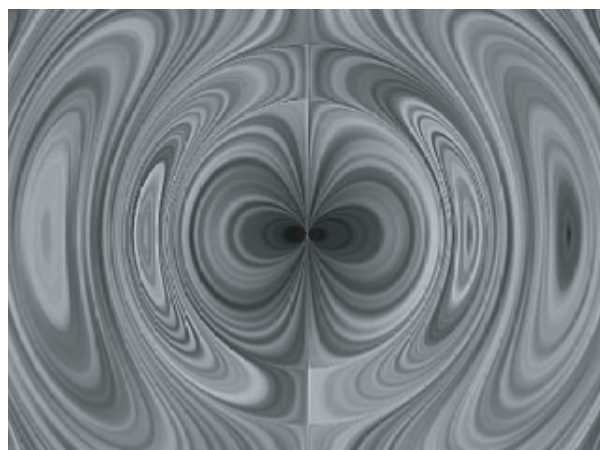
FIGURE 58 The electrical field around an oscillating *dipole* (QuickTime film © Daniel Weiskopf).

Finally, we have to integrate over all angles; this gives a factor of 2/3. In total we get

$$P = \frac{q^2 a^2}{6\pi\varepsilon_0 c^3} \ . \tag{64}$$

The total radiated power P thus depends on the square of the acceleration and on the square of the charge that is being accelerated. This is the so-called *Larmor formula*. It shows why radio transmitters need power supplies and allows deducing how large they need to be. Note that Figure 56 and Figure 57 and also show that transmitter antennas have a *preferred* direction of power emission.

Usually, electromagnetic radiation is not produced by oscillating charges, but by oscillating dipoles. A visualization of the electric field is shown in Figure 58. At large distances, a wave section can be approximated as a plane wave.

How does the world look when riding on a light beam?

At the end of the nineteenth century, the teenager Albert Einstein read a book series by Aaron Bernstein discussing the speed of light. The book asked what would happen if an observer moved at the same speed as light. Einstein thought much about the issue, and in particular, asked himself what kind of electromagnetic field he would observe in that case. Einstein later explained that this Gedanken experiment convinced him already at that young age that *nothing* could travel at the speed of light, since the field observed would have a property not found in nature. Can you find out which one he meant? Riding on a light beam situation would have strange consequences:

— You would have no mirror image, like a vampire.
— Light would not be oscillating, but would be a static field.
— Nothing would move, like in the tale of sleeping beauty.

But also at speeds *near* the velocity of light observations would be interesting. You would:
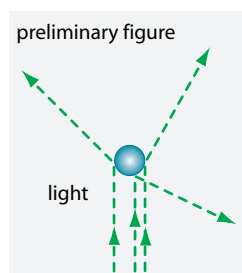
FIGURE 59 Levitating a small glass bead with a laser.

— see a lot of light coming towards you and almost no light from the sides or from behind; the sky would be blue/white in the front and red/black behind;
— observe that everything around happens very very slowly;
— experience the smallest dust particle as a deadly bullet.

Challenge 107 s   Can you think of more strange consequences? It is rather reassuring that our planet moves rather slowly through its environment, when compared to the speed of light.

Can one touch light?

If a little glass bead is put on top of a powerful laser, the bead remains suspended in mid-air, as shown in Figure 59.* This proves that light has momentum. Therefore, contrary to what we said in the beginning of our mountain ascent, images *can* be touched! In fact, the ease with which objects can be pushed even has a special name. For planets and planetoids, it is called the *albedo*, and for general objects it is called the *reflectivity*, abbreviated as r.

Like each type of electromagnetic field, and like every kind of wave, light carries en-
Challenge 108 e   ergy; the energy flow T per surface and time is

$$T = \frac{1}{\mu_0} E \times B \quad \text{giving an average} \quad \langle T \rangle = \frac{1}{2\mu_0} E_{\max} B_{\max} \ . \tag{65}$$

Obviously, light also has a momentum P. It is related to the energy E by

$$P = \frac{E}{c} \ . \tag{66}$$

Challenge 109 e   As a result, the pressure p exerted by light on a body is given by

$$p = \frac{T}{c}(1 + r) \tag{67}$$

where for black bodies we have that a reflectivity $r = 0$ and for mirrors $r = 1$; other bodies have values in between. What is your guess for the amount of pressure due to sunlight on

---

* The heaviest object that has been levitated with a laser had a mass of 20 g; the laser used was enormous, and the method also made use of a few additional effects, such as shock waves, to keep the object in the air.

FIGURE 60 The tail of comet McNaught, photographed in Australia in 2007 (© Flagstaffotos).
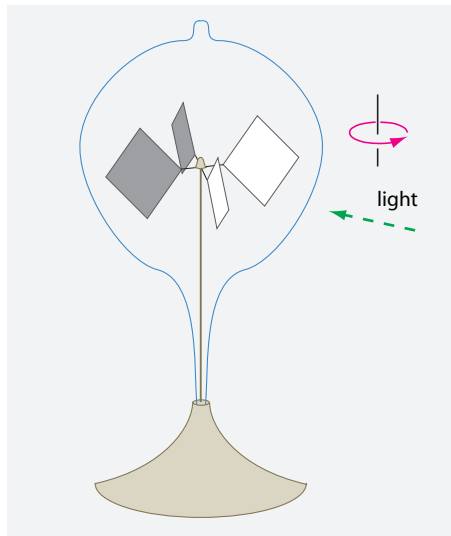
FIGURE 61 A commercial light mill turns *against* the light (Wikimedia).

Challenge 110 s a black surface of one square metre? Is this the reason that we feel more pressure during the day than during the night?

If lasers are not available, rather delicate equipment is needed to detect the momentum or the radiation pressure of light. Already in 1619, Johannes Kepler had suggested in *De cometis* that the tails of comets exist only because the light of the Sun hits the small

dust particles that detach from it. For this reason, the tail always points *away* from the
Sun, as you might want to check at the next opportunity. Today, we know that Kepler
was right; but proving the hypothesis is not easy.

In order to detect the radiation pressure of light, in 1873, William Crookes* invented
the *light mill radiometer*. The light mill consists of four thin plates, black on one side
and shiny on the other, that are mounted on a vertical axis, as shown in Figure 61. How-
ever, when Crookes finished building it – it was similar to those sold in shops today – he
found, like everybody else, that it turned in the wrong direction, namely with the shiny
side towards the light! (Why is it wrong?) You can check it by yourself by shining a laser
pointer on to it. The behaviour has been a puzzle for quite some time. Explaining it in-
volves the tiny amount of gas left over in the glass bulb and takes us too far from the
topic of our mountain ascent. It was only in 1901, with the advent of much better pumps,
that the Russian physicist Pyotr Lebedew managed to create a sufficiently good vacuum
to allow him to measure the light pressure with such an improved, true radiometer. Lebe-
dew also confirmed the predicted value of the light pressure and proved the correctness
of Kepler's hypothesis about comet tails. Today it is even possible to build tiny propellers
that start to turn when light shines on to them, in exactly the same way that the wind
turns windmills.

But light cannot only touch and be touched, it can also *grab*. In the 1980s, Arthur
Ashkin and his research group developed actual *optical tweezers* that allow one to grab,
suspend and move small transparent spheres of 1 to 20 μm diameter using laser beams. It
is possible to do this through a microscope, so that one can also observe at the same time
what is happening. This technique is now routinely used in biological research around
the world, and has been used, for example, to measure the force of single muscle fibres,
by chemically attaching their ends to glass or Teflon spheres and then pulling them apart
with such optical tweezers.

But that is not all. In the last decade of the twentieth century, several groups even
managed to *rotate* objects, thus realizing actual *optical spanners*. They are able to rotate
particles at will in one direction or the other, by changing the optical properties of the
laser beam used to trap the particle.

In fact, it does not take much to deduce that if light has linear momentum, circularly
polarized light also has *angular* momentum. In fact, for such a wave the angular momen-
tum L is given by

$$L = \frac{E_{\text{nergy}}}{\omega} \ . \tag{68}$$

Equivalently, the angular momentum of a wave is $\lambda/2\pi$ times its linear momentum. For
light, this result was already confirmed in the early twentieth century: a light beam can
put certain materials (which ones?) into rotation, as shown in Figure 62. Of course, the
whole thing works even better with a laser beam. In the 1960s, a beautiful demonstration
was performed with microwaves. A circularly polarized microwave beam from a maser
– the microwave equivalent of a laser – can put a metal piece absorbing it into rotation.
Indeed, for a beam with cylindrical symmetry, depending on the sense of rotation, the

Challenge 111 e

Challenge 112 s

Ref. 59

Ref. 60

Ref. 61

Ref. 62

Ref. 62

Challenge 113 e
Ref. 63
Challenge 114 ny

---

* William Crookes (b. 1832 London, d. 1919 London), English chemist and physicist, president of the Royal
Society, discoverer of thallium, and believer in spiritualism.
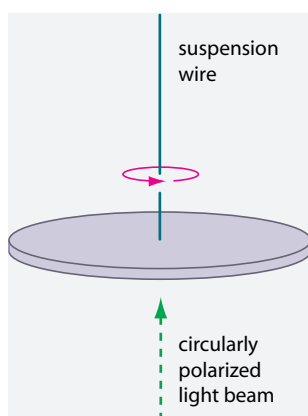
FIGURE 62 Light can rotate objects
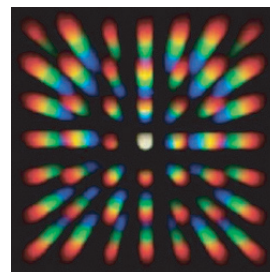
FIGURE 63 Umbrellas decompose white light: look at a small lamp through a black umbrella at night (© Wikimedia).

angular momentum is either parallel or antiparallel to the direction of propagation. All these experiments confirm that light also carries angular momentum, an effect which will play an important role in the quantum part of our mountain ascent.

We note that not for all waves angular momentum is energy per angular frequency. This is only the case for waves made of what in quantum theory will be called spin 1 particles. For example, for gravity waves the angular momentum is *twice* this value, and they are therefore expected to be made of spin 2 particles.

What does this mean for the comet tails mentioned above? The issue was settled definitely in 1986. A satellite was shot up to an altitude of 110 000 km and made to release a cloud of barium. The cloud was visible from the Earth, and it soon developed a tail that was visible from Earth: that was the first artificial comet. It turns out that comet tails shapes are partly due to hitting photons, partly due to the solar wind, and partly to magnetic fields.

In summary, light can touch and be touched. Obviously, if light can rotate bodies, it can also *be* itself rotated. Could you imagine how this can be achieved?

Ref. 64

Challenge 115 s

War, light and lies

From the tiny effects of equation (67) for light pressure we deduce that light is not an efficient tool for hitting objects. On the other hand, light is able to heat up objects, as we can feel in the sun or when the skin is touched by a laser beam of about 100 mW or more. For the same reason even cheap laser pointers are dangerous to the eye.

In the 1980s, and again in 2001, a group of people who had read too many science fiction novels managed to persuade the military – who also indulge in this habit – that lasers could be used to shoot down missiles, and that a lot of tax money should be spent on developing such lasers. Using the definition of the Poynting vector and a hitting time of about 0.1 s, are you able to estimate the weight and size of the battery necessary for such a device to work? What would happen in cloudy or rainy weather?

Challenge 116 ny
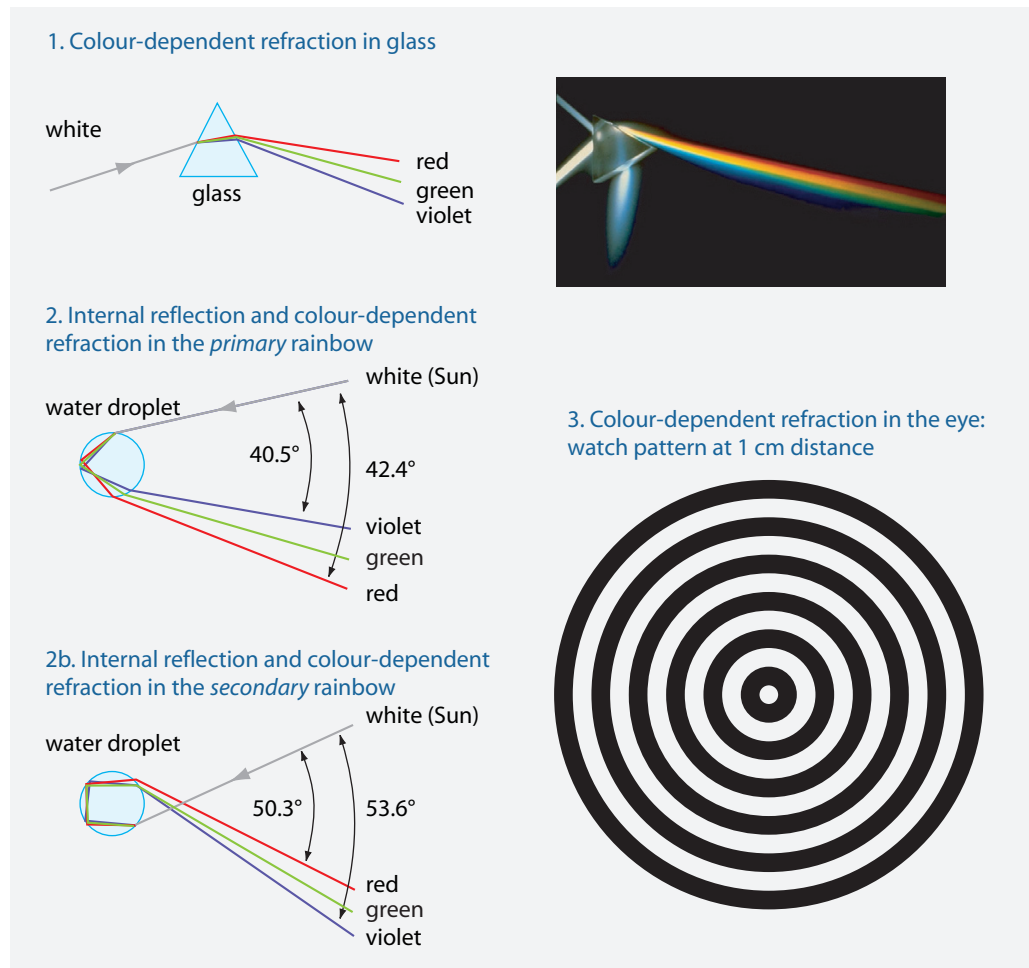
FIGURE 64 Three proofs that white light is a mixture of colours (with exaggerated angle differences): prism decomposition, rainbow formation and the coloured borders seen on a circular black and white pattern (photograph by Susan Schwartzenberg, © Exploratorium www.exploratorium.edu).

Other people tried to persuade NASA to study the possibility of propelling a rocket using emitted light instead of ejected gas. Are you able to estimate whether this is feasible?

Challenge 117 ny

What is colour?

We saw that radio waves of certain frequencies are visible. Within that range, different frequencies correspond to different colours. (Are you able to convince a friend about this?) But the story does not finish here. Numerous colours can be produced either by a single wavelength, i.e., by *monochromatic* light, or by a *mixture* of several different colours. For example, standard yellow can be, if it is pure, an electromagnetic beam of 575 nm wavelength or it can be a mixture of standard green of 546.1 nm and standard red of 700 nm. The eye cannot distinguish between the two cases; only spectrometers can. In everyday life, all colours turn out to be mixed, with the exceptions of those of yellow street lamps,

Challenge 118 s

of laser beams and of laboratory spectra. You can check this for yourself, using an umbrella or a compact disc: they decompose light mixtures, but not pure colours.

In particular, white light is a mixture of a continuous range of colours with a given intensity per wavelength. To check that white light is a mixture of colours, simply hold the lower right-hand side of Figure 64 so close to your eye that you cannot focus the stripes any more. The unsharp borders of the white stripes have a pink or a green shade. These colours are due to the imperfections of the human eye, its so-called *chromatic aberrations*. Aberrations have the consequence that not all light frequencies follow the same path through the lens of the eye, and therefore they hit the retina at different spots. This is the same effect that occurs in prisms or in water drops showing a rainbow. By the way, the shape of the rainbow tells something about the shape of the water droplets. Can you

deduce the connection?

The left-hand side of Figure 64 explains how rainbows form. The main idea is that internal reflection inside the water droplets in the sky is responsible for throwing back the light coming from the Sun, whereas the wavelength-dependent refraction at the air–water surface is responsible for the different paths of each colour. The first two persons to verify this explanation were Theodoricus Teutonicus de Vriberg (*c.* 1240 to *c.* 1318), in the

years from 1304 to 1310 and, at the same time, the Persian mathematician Kamal al-Din al-Farisi. To check the explanation, they did something smart and simple; anybody can

repeat this at home. They built an enlarged water droplet by filling a thin spherical (or cylindrical) glass container with water; then they shone a beam of white light through it. Theodoricus and al-Farisi found exactly what is shown in Figure 64. With this experiment, each of them was able to reproduce the angle of the main or primary rainbow, its colour sequence, as well as the existence of a secondary rainbow, its observed angle and

its inverted colour sequence.* All these bows are visible in Figure 49. Theodoricus's beautiful experiment is sometimes called the most important contribution of natural science in the Middle Ages.

Incidentally, the explanation of Figure 64 is not complete. It is drawn with the light ray hitting the water droplet at a specific spot. If the light ray hits the droplet at other spots, the rainbows appear at other angles; however, all those rainbows wash out. Only the visible rainbow remains, because its deflection angles are extremal.

Incidentally, at sunset the atmosphere itself also acts as a prism, or more precisely, as a cylindrical lens affected by spherochromatism. That means that the Sun is split into different images, one for each colour, which are slightly shifted with respect to each other;

the total shift is about 1% of the diameter. As a result, the rim of the Sun is coloured. If the weather is favourable, if the air is clear up to and beyond the horizon, and if the correct temperature profile is present in the atmosphere, a colour-dependent mirage will appear. As a result, for about a second it will be possible to see, after or near the red, orange and yellow images of the setting Sun, the green–blue image, sometimes even detached. This

is the famous *green flash* described by Jules Verne in his novel *Le Rayon-vert*. It is often

---

* Can you guess where the tertiary and quaternary rainbows are to be seen? There are rare reported sightings of them. The hunt to observe the fifth-order rainbow is still open. (In the laboratory, bows around

droplets up to the thirteenth order have been observed.) For more details, see the beautiful website at www. atoptics.co.uk. There are several formulae for the angles of the various orders of rainbows; they follow from straightforward geometric considerations, but are too involved to be given here.
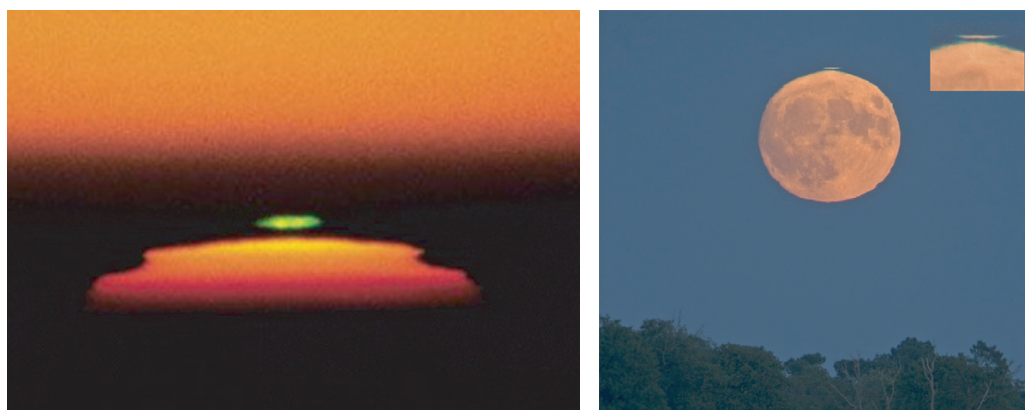
FIGURE 65 A green flash above the setting Sun and one above the Moon, showing also the colour change of the Moon rim (© Andrew Young and Laurent Laveder/PixHeaven.net).

FIGURE 66 Milk and water simulate the evening sky (© Antonio Martos).

seen on tropical beaches, for example in Hawaii, and from the decks of ships in warm waters.

Even pure air splits white light. However, this is not done by dispersion, but by scattering. Wavelength-dependent scattering is the reason that the sky and far away mountains look blue or that the Sun looks red at sunset and at sunrise. (The sky looks black even during the day from the Moon.) You can repeat this effect by looking through water at a black surface or at a lamp. Adding a few drops of milk to the water makes the lamp yellow and then red, and makes the black surface blue (like the sky seen from the Earth as compared to the sky seen from the Moon) as shown in Figure 66. More milk increases the effect. For the same reason, sunsets are especially red after volcanic eruptions.

To clarify the difference between colours in physics and colour in human perception and language, a famous linguistic discovery deserves to be mentioned: colours in human language have a natural *order*. Colours are ordered by all peoples of the world, whether they come from the sea, the desert or the mountains, in the following order: 1. black

and white, 2. red, 3. green and yellow, 4. blue, 5. brown, 6. mauve, pink, orange, grey and sometimes a twelfth term that differs from language to language. (Colours that refer to objects, such as aubergine or sepia, or colours that are not generally applicable, such as blond, are excluded in this discussion.) The precise discovery is the following: if a particular language has a word for any of these colours, then it also has a word for all the preceding ones. The result also implies that people use these basic colour classes even if their language does *not* have a word for each of them. These strong statements have been confirmed for over 100 languages.

Fun with rainbows

The width of the usual, primary rainbow is 2.25°, for the secondary rainbow it is about twice that value (which is one reason why it is less bright). The width is larger than the dispersion angle difference given in Figure 64 because the angular size of the sun, about 0.5°, has (roughly) to be added on top of the angle difference.

If the droplets are very fine, the rainbow become white; it is then called a *fogbow*. Such bows are also often seen from aeroplanes. If the droplets are not round, for example due to strong wind, one can get a so-called *irregular* or *twinned rainbow*. An example is shown in Figure 67.

Light from the rainbow is tangentially polarized. You can check that easily with polarizing sunglasses. During the internal reflection in the water droplets, as the reflection angle is very near to the angle at which total reflection sets in, light gets polarized. (Why is does this lead to polarization?) More on polarization will be told in the next section.

If the air is full of ice crystals instead of droplets, the situation changes again. One can then get additional images of the sun in the direction of the sun. They are called *parhelia* (or sundogs). This happens most clearly with no wind, if the crystals are all oriented in the same direction. In that case one can take photographs such as the one shown in Figure 68.

What is the speed of light? – Again

Physics talks about motion. Talking is the exchange of sound; and sound is an example of a signal. A *(physical) signal* is the transport of information using the transport of energy. There are no signals without a motion of energy. Indeed, there is no way to store information without storing energy. To any signal we can thus ascribe a propagation speed. The highest possible signal speed is also the maximal velocity of the general influences, or, to use sloppy language, the maximal velocity with which effects spread causes.

If the signal is carried by matter, such as by the written text in a letter, the signal velocity is then the velocity of the material carrier, and experiments show that it is limited by the speed of light.

For a wave carrier, such as water waves, sound, light or radio waves, the situation is less evident. What is the speed of a wave? The first answer that comes to mind is the speed with which wave crests of a sine wave move. This already introduced *phase velocity* is given by the ratio between the frequency and the wavelength of a monochromatic wave, i.e., by

$$v_{\text{ph}} = \frac{\omega}{k} \; . \tag{69}$$

FIGURE 67 Five rare types of rainbows: a fogbow (top left), an irregular, split rainbow in a windy situation due to non-spherical rain drops (top right, shown with increased colour saturation), a six-fold rainbow (middle left), a red rainbow at sunset (middle right), and a moonbow, created by the Moon, not by the Sun, and brightened digitally (© Michel Tournay, Eva Seidenfaden, Terje Nordvik, Zhu XiaoJin and Laurent Laveder).

For example, the phase velocity determines interference phenomena. Light in a vacuum has the same phase velocity $v_{ph} = c$ for all frequencies. Are you able to imagine an exper-

iment to test this to high precision?

On the other hand, there are cases where the phase velocity is greater than c, most notably when light travels through an absorbing substance, and when at the same time the frequency is near to an absorption maximum. In these cases, experiments show that

the phase velocity is *not* the signal velocity. For such situations, a better approximation

FIGURE 68 A composite photograph showing the parhelia, the light pillars, the halo and the upper tangent arc formed by ice crystals in the air, if all oriented in the same direction (© Phil Appleton).
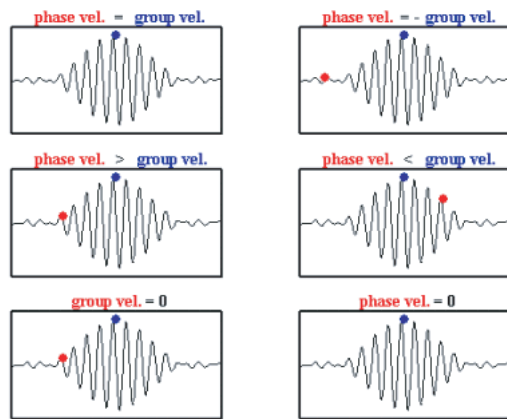
FIGURE 69 A visualisation of group velocity (blue) and phase velocity (red) for different types of waves (QuickTime film © ISVR, University of Southhampton).
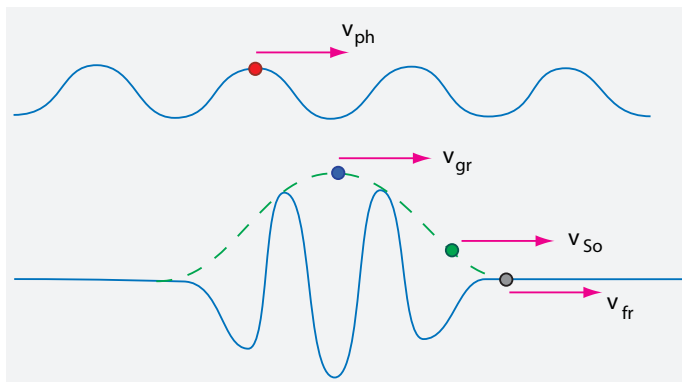
FIGURE 70 The definition of the important velocities in wave phenomena, including Sommerfeld's front velocity and the forerunner velocity.

to the signal speed is the *group velocity*, i.e., the velocity at which a group maximum will

travel. This velocity is given by

$$v_{\mathrm{gr}} = \left.\frac{\mathrm{d}\omega}{\mathrm{d}k}\right|_{k_0} \; , \qquad\qquad\qquad (70)$$

where k_0 is the central wavelength of the wave packet. We observe that $\omega = c(k)k = 2\pi v_{\mathrm{ph}}/\lambda$ implies the relation

$$v_{\mathrm{gr}} = \left.\frac{\mathrm{d}\omega}{\mathrm{d}k}\right|_{k_0} = v_{\mathrm{ph}} - \lambda\frac{\mathrm{d}v_{\mathrm{ph}}}{\mathrm{d}\lambda} \; . \qquad\qquad (71)$$

This means that the sign of the last term determines whether the group velocity is larger or smaller than the phase velocity. For a travelling group, as shown by the dashed line in Figure 70, this means that new maxima appear either at the end or at the front of the group. Experiments show that this is only the case for light passing *through matter*; for light *in vacuum*, the group velocity has the same value $v_{\mathrm{gr}} = c$ for all values of the wave vector k.

You should be warned that many publications are still propagating the incorrect statement that the group velocity *in a material* is never greater than c, the speed of light in vacuum. Actually, the group velocity in a material can be zero, infinite or even negative; this happens when the light pulse is very narrow, i.e., when it includes a wide range of frequencies, or again when the frequency is near an absorption transition. In many (but not all) cases the group is found to widen substantially or even to split, making it difficult to define precisely the group maximum and thus its velocity. Many experiments have confirmed these predictions. For example, the group velocity in certain materials has been measured to be *ten times* that of light. The refractive index then is smaller than 1. However, in all these cases the group velocity is *not* the same as the signal speed.[*]

What then is the best velocity describing signal propagation? The German physicist Arnold Sommerfeld[**] almost solved the main problem in the beginning of the twentieth century. He defined the signal velocity as the velocity v_{So} of the front slope of the pulse, as shown in Figure 70. The definition cannot be summarized in a formula, but it does have the property that it describes signal propagation for almost all experiments, in particular those in which the group and phase velocity are larger than the speed of light. When studying its properties, it was found that for no material is Sommerfeld's signal velocity greater than the speed of light in vacuum.

Sometimes it is conceptually easier to describe signal propagation with the help of the energy velocity. As previously mentioned, every signal transports energy. The *energy velocity* v_{en} is defined as the ratio between the energy flow density S, i.e., the Poynting

---

[*] In quantum mechanics, Schrödinger proved that the velocity of an electron is given by the group velocity of its wave function. Therefore the same discussion reappeared in quantum theory, as we will find out in the next volume of our mountain ascent.

[**] Arnold Sommerfeld (b. 1868 Königsberg, d. 1951 München) was a central figure in the spread of special and general relativity, of quantum theory, and of their applications. A professor in Munich, an excellent teacher and text book writer, he worked on atomic theory, on the theory of metals and on electrodynamics, and was the first to understand the importance and the mystery around 'Sommerfeld's famous fine structure constant.'

vector, and the energy density W, both taken in the direction of propagation. For electromagnetic fields – the only ones fast enough to be interesting for eventual superluminal signals – this ratio is

$$v_{\mathrm{en}} = \frac{\langle P \rangle}{\langle W \rangle} \; . \tag{72}$$

However, as in the case of the front velocity, in the case of the energy velocity we have to specify the underlying averaging procedure, denoted by $\langle \rangle$, i.e., whether we mean the energy transported by the main pulse or by the front of it. In vacuum, neither is ever greater than the speed of light.* (In general, the velocity of energy in matter has a value slightly different from Sommerfeld's signal velocity.)

Ref. 71

In recent years, the progress in light detector technology, allowing one to detect even the tiniest energies, has forced scientists to take the fastest of all these energy velocities to describe signal velocity. Using detectors with the highest possible sensitivity we can use as signal the first point of the wave train whose amplitude is different from zero, i.e., the first tiny amount of energy arriving. This point's velocity, conceptually similar to Sommerfeld's signal velocity, is commonly called the *front velocity* or, to distinguish it even more clearly from Sommerfeld's case, the *forerunner velocity*. It is simply given by

Challenge 127 s

$$v_{\mathrm{fr}} = \lim_{\omega \to \infty} \frac{\omega}{k} \; . \tag{73}$$

The forerunner velocity is *never* greater than the speed of light in a vacuum, even in materials. In fact it is precisely c because, for extremely high frequencies, the ratio ω/k is independent of the material, and vacuum properties take over. The forerunner velocity is the true signal velocity or the *true velocity of light*. Using it, all discussions on light speed become clear and unambiguous.

To end this section, here are two challenges for you. Which of all the velocities of light is measured in experiments determining the velocity of light, e.g. when light is sent to the Moon and reflected back? And now a more difficult one: why is the signal speed of light inside matter less than the speed in vacuum, as all experiments show?

Challenge 128 s

Challenge 129 s

Signals and predictions

When one person reads a text over the phone to a neighbour who listens to it and maybe repeats it, we speak of communication. For any third person, the speed of communication is always less than the speed of light. But if the neighbour already knows the text, he can recite it without having heard the readers' voice. To the third observer such a situation appears to imply motion that is faster than light. Prediction can thus *mimic* communication and, in particular, it can mimic faster-than-light (superluminal) communication. Such a situation was demonstrated most spectacularly in 1994 by Günter Nimtz, who seemingly transported music – all music is predictable for short time scales

Ref. 75

---

* Signals not only carry energy, they also carry negative entropy ('information'). The entropy of a transmitter increases during transmission. The receiver decreases in entropy (but less than the increase at the transmitter, of course).

Ref. 73

Note that the negative group velocity implies energy transport against the propagation velocity of light. This is possible only in *energy loaded* materials.

Ref. 74

TABLE 14 Experimental properties of flat vacuum (and of the 'aether').

Physical property	Experimental value
permeability	μ_0 =1.3 µH/m
permittivity	ε_0 =8.9 pF/m
wave impedance/resistance	Z_0 = 376.7 Ω
conformal invariance	applies
spatial dimensionality	3
topology	R^3
mass and energy content	none
friction on moving bodies	none
motion	none

– through a 'faster-than-light' system. To distinguish between the two situations, we note that in the case of prediction, no transport of energy takes place, in contrast to the case of communication. In other words, the definition of a signal as a transporter of information is not as useful and clear-cut as the definition of a signal as a *transporter of energy*. In the above-mentioned experiment, no energy was transported faster than light. The same distinction between prediction on the one hand and signal or energy propagation on the other will be used later to clarify some famous experiments in quantum mechanics.

> " If the rate at which physics papers are being published continues to increase, physics journals will soon be filling library shelves faster than the speed of light. This does not violate relativity since no useful information is being transmitted. "
>
> David Mermin

Aether good-bye

Gamma rays, X-rays, light and radio waves are moving electromagnetic waves. All exist in empty space. What is oscillating when light travels? Maxwell himself called the oscillating 'medium' the *aether*. The properties of the aether measured in experiments are listed in Table 14. The strange numerical values are due to the definition of the units henry and farad.

The last item of Table 14 is the most important: despite intensive efforts, nobody has been able to detect any *motion* of the aether. In other words, even though the aether supposedly oscillates, it does not move. Together with the other data, all these results can be summed up in one sentence: there is no way to distinguish the aether from the vacuum: they are one and the same.

Sometimes one hears that certain experiments or even the theory of relativity show that the aether does not exist. This is not strictly correct. In fact, experiments show something more important. All the data show that *the aether is indistinguishable from the vacuum*. Of course, if we use the change of curvature as the definition for motion of the

vacuum, the vacuum *can* move, as we found out in the section on general relativity; but the aether still remains indistinguishable from it.*

Later we will even find out that the ability of the vacuum to allow the propagation of light and its ability to allow the propagation of particles are equivalent: both require the same properties. Therefore the aether remains *indistinguishable* from a vacuum in the rest of our walk. In other words, the aether is a superfluous concept; we will drop it from our walk from now on. Despite this result, we have not yet finished the study of the vacuum; vacuum will keep us busy for a long time, starting with the following part of our ascent. Moreover, quite a few of the aspects in Table 14 will require some amendments later.

Ref. 77

Challenges and fun curiosities about light

Since light is a wave, something must happen if it is directed to a hole less than its wavelength in diameter. What exactly happens?

Challenge 131 s

* *

Electrodynamics shows that light beams always push; they never pull. Can you confirm that 'tractor beams' are impossible in nature?

Challenge 132 e

* *

It is well known that the glowing material in light bulbs is tungsten wire in an inert gas. This was the result of a series of experiments that began with the grandmother of all lamps, namely the cucumber. The older generation knows that a pickled cucumber, when attached to the 230 V of the mains, glows with a bright green light. (Be careful; the experiment is dirty and dangerous.)

* *

If the light emitted by the headlights of cars were polarized from the bottom left to the upper right (as seen from the car's driver) one could vastly improve the quality of driving at night: one could add a polarizer to the wind shield oriented in the same direction. As a result, a driver would see the reflection of his own light, but the light from cars coming towards him would be considerably dampened. Why is this not done in modern cars?

Challenge 133 ny

Summary on light

In summary, radio waves, infrared light, visible light, ultraviolet light, X-rays and gamma rays are electromagnetic waves whose dispersion relation in vacuum is $\omega = ck$, where $c = 299\,792\,458$ m/s. Electromagnetic waves carry energy, linear momentum and angular momentum, and move faster than any material object. The speed of electromagnetic waves c is the limit energy speed in nature.

Ref. 77   * In fact, the term 'aether' has been used as an expression for several different ideas, depending on the author. First of all it was used for the idea that a vacuum is not empty, but *full*; secondly, that this fullness can be described by *mechanical models*, such as gears, little spheres, vortices, etc.; thirdly, it was imagined that a vacuum is *similar to matter*, being made of the same substrate. Interestingly, these issues will reappear in the last part of our mountain ascent.
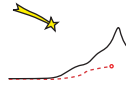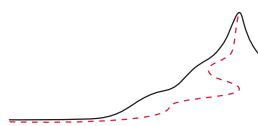
IMAGES AND THE EYE – OPTICS

O PTICS is the field that explores the production of images. In particular, ptics is the study and use of light *production*, of light *transport*, and f light and image *detection*. With this definition of optics, we note directly that classical electrodynamics can describe only the transport of light. The production and the detection of light are always quantum effects. Every lamp is a device based on quantum physics. Every detector of light, including the eye, is based on quantum physics. Therefore, in this chapter we mainly explore the motion of light and the way it forms images, and give only a short introduction into light sources and the eye.

Ways to produce images

Producing images is an important part of modern society. The quality of images depends on the smart use of optics, electronics, computers and materials science. Despite the long history of optics, there are still new results in the field. Images, i.e., two or three-dimensional reproductions, can be taken by at least six groups of techniques:

— *Photography* uses a light source, lenses and film – or another large area detector. Photography can be used in reflection, in transmission, with phase-dependence, with various illuminations, and with light sources and detectors for various wavelengths.
— *Optical microscopy* uses a light source, lenses and film (or another large area detector). If the illumination is through the sample, in transmission, one speaks of *bright-field microscopy*. (Variations use coloured or polarizing filters.) If the illumination is from the side, one speaks of *oblique microscopy*. If the illumination is confined to an outer ring of light, one speaks of *dark-field microscopy*. An even more elaborate illumination system, using plane waves, allows *phase-contrast microscopy*. (It was invented by Frits Zernike in the 1930s and earned him the Nobel Prize in Physics in 1953.) If one splits a polarized illumination beam into two components that pass the sample at close (but not identical) locations, and then recombines them afterwards, one speaks of *differential interference contrast microscopy*. If a sample is treated with a fluorescent dye, the illuminating light is filtered out, and only the fluorescence is observed, one speaks of *fluorescence microscopy*. The image quality of expensive microscopes can be further improved with the help of a computer, with the help of deconvolution techniques.
— *Telescopy* is used most of all in geodesy and astronomy. The most advanced astronomical telescopes can compensate star images for the effects of the turbulence of the atmosphere; they can also take images at various wavelengths, ranging from ra-

FIGURE 71 An X-ray photographic image of a ten-year old boy with polydactyly (courtesy Drgnu23/Wikimedia).

dio frequencies, infrared, visible, ultraviolet to X-rays. X-ray telescopes have to be operated outside the atmosphere, to avoid absorption by air, for example on rockets, satellites or high-altitude balloons.

— *Holography* uses lasers and large area detectors. Holography allows taking three-dimensional images of objects. It is usually used in reflection, but can also be used in transmission.

— *Tomography*, usually performed in transmission, uses a source and a line detector that are rotated together around an object. This technique allows imaging cross sections of materials. For example, light tomography is a promising technique, without any health risk, for breast cancer detection.

— *Scanning techniques* construct images point by point through the motion of the detector, the light source or both. There are numerous scanning microscopy techniques: *confocal laser scanning microscopy*, the fibre-based *near-field scanning optical microscopy*, and combinations of them with fluorescence techniques or various deconvolution techniques. Many of these scanning microscopy techniques allow resolutions much lower than the wavelength of light, which is impossible with usual optical microscopy.

Each imaging method can be used with radio waves, infrared, visible light, ultraviolet light, X-rays or gamma rays. In all imaging methods, the race is for images with the highest resolution possible. But let us start with light sources.

LIGHT SOURCES

Without light sources, there would be no images. The most important light sources of visible and infrared light are black bodies. Let us see why.
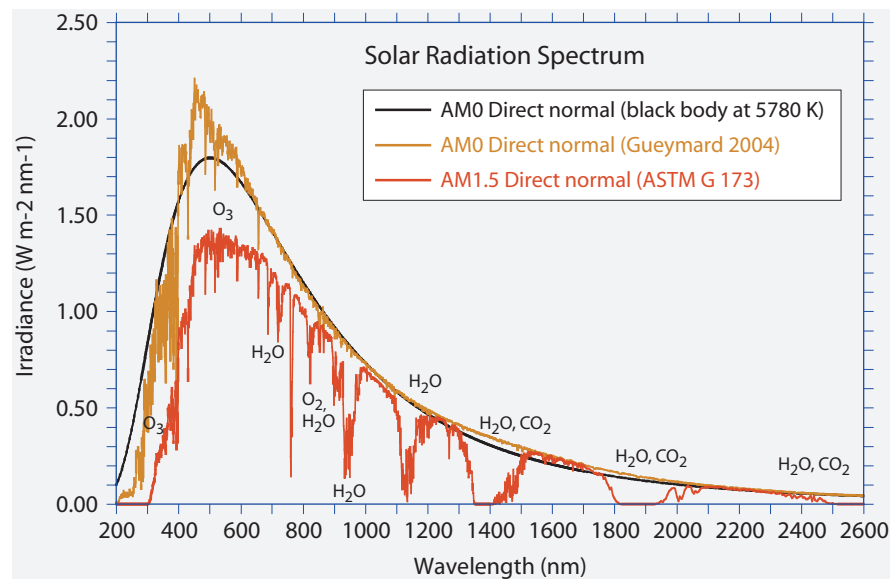
FIGURE 72 A black body spectrum at 5780 K, the solar spectrum *above* the atmosphere in direction of the Sun, with 1350 W/m², and the spectrum with 1.5 air masses in between, with 844 W/m², giving roughly the spectrum of a typical sunny day at sea level; the gases responsible for the absorption bands are also shown (© Chris Gueymard).

WHY CAN WE SEE EACH OTHER? BLACK BODIES AND THE TEMPERATURE OF LIGHT

Physicists have a strange use of the term 'black'. A body that glows perfectly is called a *black body*. In this domain, 'perfect' means that the surface of the body has *no* effect on its colour.

> ▷ A **black body** is a body that absorbs all radiation impinging on it.

In other words, a black body is a body without reflection or transmission of radiation. Black bodies are an idealization; above all, they are only black at low temperature. At high temperature, they glow or shine in brown, red, orange, yellow, white or blue. The essence of black bodies is that the colour they have, i.e., the light they radiate, is independent of the surface. Black bodies are thus ideal in this sense. Real bodies, which do show surface effects, can be classified by their *emissivity*. The emissivity gives the degree to which a body approaches a black body. Mirrors have emissivities of around 0.02, whereas black soot can have values as high as 0.95. Practically all bodies at everyday temperature are not black bodies: their colour is not determined by emission, but mostly by the absorption and reflection of light at their surface.

Black bodies, as the section on quantum theory will show, have *smooth* light emission spectra. An example for a spectrum of a black body, and for a spectrum of a real body – in this case, the Sun – is shown in Figure 72.

Black bodies are also used to define the colour *white*. What we commonly call *pure white* is the colour emitted by the Sun. The sun is not a good black body, as Figure 72

shows (its effective temperature is 5780 K). Because of these problems, pure white is now
defined as the colour of a black body of 6500 K, e.g. by the Commission Internationale
d'Eclairage. Hotter black bodies are bluish, colder ones are yellow, orange, red, brown or
black.The stars in the sky are classified in this way.

Black bodies are thus bodies that glow perfectly. Most real bodies are only rough ap-
proximations of black bodies, even at temperatures at which they shine yellow light. For
example, the tungsten in incandescent light bulbs, at around 2000 K, has an emissivity of
around 0.4 for most wavelengths, so that its spectrum is a corresponding fraction of that
of black body. (However, the glass of the light bulb then absorbs much of the ultraviolet
and infrared components, so that the final spectrum is not at all that of a black body.)

Black body radiation has two important properties: first, the emitted light power in-
creases with the fourth power of the temperature. With this relation alone you can check
the temperature of the Sun, mentioned above, simply by comparing the size of the Sun
with the width of your thumb when your arm is stretched out in front of you. Are you
able to do this? (Hint: use the excellent approximation that the Earth's average tempera-
ture of about 14.0°C is due to the Sun's irradiation.)

The precise expression for the emitted energy density u per frequency v can be de-
duced from the radiation law for black bodies discovered by Max Planck*

$$u(v, T) = \frac{8\pi h}{c^3} \frac{v^3}{e^{hv/kT} - 1} \ . \tag{74}$$

He made this important discovery, which we will discuss in more detail in the quantum
part of our mountain ascent, simply by comparing this curve with experiment. The new
constant h is called the *quantum of action* or *Planck's constant* and turns out to have the
value $6.6 \cdot 10^{-34}$ Js, and is central to all quantum theory, as we will see. The other constant
Planck introduced, the Boltzmann constant k, appears as a prefactor of temperature all
over thermodynamics, as it acts as a conversion unit from temperature to energy.

The radiation law gives for the total emitted energy density the expression

$$u(T) = T^4 \ \frac{8\pi^5 k^4}{15c^3 h^3} \tag{75}$$

from which equation (81) is deduced using $I = uc/4$. (Why?)

The second property of black body radiation is the value of the peak wavelength,
i.e., the wavelength emitted with the highest intensity. This wavelength determines their

---

* Max Planck (1858–1947), professor of physics in Berlin, was a central figure in thermostatics. He discov-
ered and named *Boltzmann's constant k* and the *quantum of action h*, often called Planck's constant. His
introduction of the quantum hypothesis gave birth to quantum theory. He also made the works of Einstein
known in the physical community, and later organized a job for him in Berlin. He received the Nobel Prize
for physics in 1918. He was an important figure in the German scientific establishment; he also was one of
the very few who had the courage to tell Adolf Hitler *face to face* that it was a bad idea to fire Jewish pro-
fessors. (He got an outburst of anger as answer.) Famously modest, with many tragedies in his personal life,
he was esteemed by everybody who knew him.

colour; it is deduced from equation (74) to be

$$\lambda_{\max} = \frac{1}{T}\frac{hc}{4.956\,k} = \frac{2.90\,\mathrm{mm\,K}}{T} \quad \text{but} \quad \hbar\nu_{\max} = T\cdot 2.82\,k/h = T\cdot 5.9\cdot 10^{10}\,\mathrm{Hz/K}\ . \quad (76)$$

Either of these expressions is called *Wien's colour displacement rule* after its discoverer.[*] The colour change with temperature is used in optical thermometers; this is also the way the temperatures of stars are measured. For 37°C, human body temperature, it gives a peak wavelength of 9.3 μm or 115 THz, which is therefore the colour of the bulk of the radiation emitted by every human being. (The peak wavelength does not correspond to the peak frequency. Why?) On the other hand, following the telecommunication laws of many countries, any radiation emitter needs a licence to operate; it follows that strictly in Germany only dead people are legal, and only if their bodies are at absolute zero temperature.

We saw that a black body – or a star – can be blue, white, yellow, orange, red or brown. A black body is never green. Can you explain why?

Above, we predicted that any material made of charges emits radiation. Are you able to find a simple argument showing whether heat radiation is or is not this classically predicted radiation?

But let us come back to the question in the section title. The existence of thermal radiation implies that any hot body will cool, even if it is left in the most insulating medium there is, namely in vacuum. More precisely, if the vacuum is surrounded by a wall, the temperature of a body in the vacuum will gradually approach that of the wall.

Interestingly, when the temperature of the wall and of the body inside have become the same, something strange happens. The effect is difficult to check at home, but impres- sive photographs exist in the literature.

One arrangement in which walls and the objects inside them are at the same temperature is an *oven*. It turns out that it is *impossible* to see objects in an oven using the light coming from thermal radiation. For example, if an oven and all its contents are red hot, taking a picture of the inside of the oven (without a flash!) does not reveal anything; no contrast nor brightness changes exist that allow one to distinguish the objects from the walls or their surroundings. Can you explain the finding?

In short, we are able to see each other only because the light sources we use are at a *different* temperature from us. We can see each other only because we do *not* live in thermal equilibrium with our environment.

Limits to the concentration of light

Light sources should be as bright as possible. Are there any limits? Interestingly, for black body radiation there is an important and instructive limitation.

If we build a large lens or a curved mirror, we can collect the light of the Sun and focus it on a very small spot. Everybody has used a converging lens as a child to burn black

---

[*] Wilhelm Wien (b. 1864 Gaffken, d. 1928 München), East-Prussian physicist; he received the Nobel Prize for physics in 1911 for the discovery of this relation.

Note that the value appearing in Wien's rule can be uniquely calculated from equation (74), but cannot be expressed as a formula. Indeed, Wien's constant contains the solution of the equation $x = 5(1 - e^{-x})$.
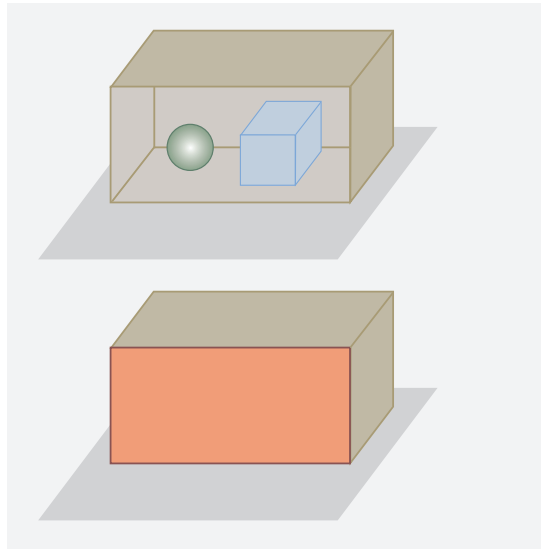
FIGURE 73 Bodies inside an oven at room temperature differ in colour, in contrast to bodies at high temperature (photo © Wolfgang Rueckner).

FIGURE 74 The last mirror of the solar furnace at Odeillo, in the French Pyrenees (© Gerhard Weinrebe).

spots on newspapers – or ants – in this way. In Odeillo, in Spain, wealthier researchers have built a curved mirror as large as a house, in order to study solar energy use and material behaviour at high temperature. Essentially, the mirror provides a cheap way to fire an oven in its focus. (And 'focus' is the Latin word for 'hearth'.)

Kids find out quite rapidly that large lenses allow them to burn things more easily than small ones. It is obvious that the Odeillo site is the record holder in this game. However, building a larger mirror does not make much sense. Whatever its size may be, such a set-up *cannot* reach a higher temperature than that of the original light source. The surface temperature of the Sun is about 5780 K; indeed, the highest temperature reached so far is about 4000 K. Are you able to show that this limitation is equivalent to the second law of thermodynamics, as Hemholtz, Clausis and Airy showed?

In short, nature provides a *limit* to the concentration of light energy. More precisely, we can say: thermodynamics limits heating with thermal light sources. In fact, we will encounter additional limits in the course of our exploration.

Ref. 81

Challenge 142 ny

MEASURING LIGHT INTENSITY

Light sources differ in brightness. Measuring what we call 'dark' and 'bright' is somewhat involved, because light can be diffuse or directed. To achieve proper measurements, the SI, the international system of units, defines a specific base unit, the candela:

> ■ 'The *candela* is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency $540 \cdot 10^{12}$ hertz and has a radiant intensity in that direction of (1/683) watt per steradian.'

The candela is thus a unit for light power per (solid) angle, usually called *luminous intensity*, except that it is corrected for the eye's sensitivity: the candela measures only *visible* power per angle. The definition of the candela simply says that 683 cd = 683 lm/sr corresponds to 1 W/sr.

Total light power, irrespective of its direction, is measured in lumen. Therefore, 683 lm = 683 cd sr corresponds to 1 W. In other words, both the lumen and the watt measure power, or energy flux, but the lumen measures only the *visible* part of the power or energy flux. This difference is expressed by adding 'luminous' or 'radiant': thus, the lumen measures *luminous* flux, whereas the Watt measures *radiant* flux.

The factor 683 is historical. An ordinary candle emits a luminous intensity of about a candela. To put this into perspective: at night, a candle can be seen up to a distance of 10 or 20 kilometres. A 100 W incandescent light bulb produces 1700 lm, and the brightest commercial light emitting diodes about 20 lm, though laboratory devices exceed 1000 lm. Cinema projectors produce around 2 Mlm, and the brightest flashes, like lightning, 100 Mlm.

The *irradiance* of sunlight is about 1300 W/m^2 on a sunny day; on the other hand, the *illuminance* is only 120 klm/m^2 = 120 klx or 170 W/m^2. A cloud-covered summer day or a clear winter day produces about 10 klx. These numbers show that most of the energy from the Sun that reaches the Earth is outside the visible spectrum.

Illuminance is essentially what we call 'brightness' in everyday life. On a glacier, near the sea shore, on the top of a mountain, or in particular weather condition the brightness can reach 150 klx. Museums are often kept dark because water-based paintings are degraded by light above 100 lx, and oil paintings by light above 200 lx. The eyes lose their ability to distinguish colours somewhere between 0.1 lx and 0.01 lx; the eye stops to work below 1 nlx. Technical devices to produce images in the dark, such as night goggles, start to work at 1 µlx. By the way, the human body itself *shines* with about 1 plx, a value too small to be detected with the eye, but easily measured with specialized apparatus. The origin of this emission is still a topic of research.

The highest achieved light intensities, produced with high-power lasers, are in excess of 10^{18} W/m^2, more than 15 orders of magnitude higher than the intensity of sunlight. (How much is that in lux?) Such intensities are produced by tight focusing of pulsed laser beams. The electric field in such light pulses is of the same order as the field inside atoms; such a laser beam therefore ionizes all matter it encounters, including the air.

The *luminous density* is a quantity often used by light technicians. Its unit is 1 cd/m^2, unofficially called 1 Nit and abbreviated 1 nt. Human eyes see using rods only from 0.1 µcd/m^2 to 1 mcd/m^2; they see with cones only above 5 cd/m^2. Eyes see best between 100 and 50 000 cd/m^2, and they get completely overloaded above 10 Mcd/m^2: a total

TABLE 15 Some measured illuminance values.

OBSERVATION	ILLUMINANCE
Brightness of the human body	1 plx
Faint star	0.1 nlx
Sirius	10 μlx
phot (old illuminance unit)	10 μlx
Jupiter	20 μlx
Dark, moonless night	1 mlx
Full moon	0.01 to 0.24 lx
Street at night, low traffic, poor lighting	0.1 to 3 lx
Street at night, high traffic	10 to 30 lx
For reading	50 to 100 lx
Cinema screen	100 lx
Workplace	0.2 to 5 klx
Cloudy day	1 klx
Brightest lamps, used for surgery	120 klx
Sunny day	120 klx
Film in cinema projector	5 Mlx
Painful to the eye	100 Mlx

range of 15 orders of magnitude. Very few technical detectors achieve this range.

Other light and radiation sources

Apart from black bodies, many other types of light sources exist, from glowing fish to high-power lasers. They range in size from an atom to a building, in cost up to hundreds of millions of Euros, and in lifetime from a fraction of a second to hundreds of years. Most sources of visible, infrared and ultraviolet light will be covered later in our

adventure.

Sources of radio waves are common; mobile phones, radio transmitters, walkie-talkies are all sources of radio waves. They are rarely used for imaging, with one important ex-

ception. Since many stars are radio emitters, one can image the sky at these wavelengths.

On the other end of the spectrum, light sources that emit X-rays and gamma rays are common. They are routinely used in medicine and material science. Such sources are dangerous to humans, so that special care has to be taken when using them.

Radiation as weapon?

High-intensity radiation is dangerous. In many countries, there is more money to study assault weapons than to increase the education and wealth of their citizen. Several types of assault weapons using electromagnetic radiation are being researched. Two are particularly advanced.

The first weapon is a truck with a movable parabolic antenna on its roof, about 1 m in

size, that emits a high power – a few kW – microwave beam at 95 GHz. The beam, like all microwave beams, is invisible; depending on power and beam shape, it is painful or lethal, up to a distance of a 100 m. This terrible device, with which the operator can make many victims even by mistake, was ready in 2006. (Who expects that a parabolic antenna is dangerous?) Efforts to ban it across the world are slowly gathering momentum.

The second weapon under development is the so-called *pulsed impulse kill laser*. The idea is to take a laser that emits radiation that is not absorbed by air, steam or similar obstacles. An example is a pulsed deuterium fluoride laser that emits at 3.5 μm. This laser burns every material it hits; in addition, the evaporation of the plasma produced by the burn produces a strong hit, so that people hit by such a laser are hurt and hit at the same time. Fortunately, it is still difficult to make such a device rugged enough for practical mobile use. Nevertheless, experts expect battle lasers to appear soon. In short, it is probable that radiation weapons will appear in the coming years. What the men working on such developments tell their children when they come home in the evening is not clear, though.

IMAGES – TRANSPORTING LIGHT

All images are formed by transporting light in a useful manner along known paths. The simplest path is the straight line.

Making images with mirrors

Since light moves in a straight line, a flat mirror produces an image of the same size than the original. Curved mirrors can be used to increase, reduce and distort images. Expensive room mirrors are often slightly curved, in order to make people appear thinner.

Most human-made mirrors are made of metal; but living systems cannot produce pure metals. On the other hand, in living systems, mirrors abound: they are found as the tapetum in the eyes, on fish scales, on bugs, etc. How does nature produce mirrors, despite lacking the ability to use pure metals? It turns out that sandwiches of different thin transparent materials – one of which is typically crystalline guanine – can produce mirrors that are almost as good as metal mirrors. Such mirrors are based on interference effects and are called *dielectric mirrors*.

Image-forming mirrors are used in large telescopes, in systems for X-rays, in devices by medical doctors, and, if one wants to count this, for the shaping of light beams in cars or in pocket lamps. Interestingly, also some living beings use mirrors for imaging. The most famous example is the spookfish shown in Figure 75. It is able to look up and down at the same time, and does so using mirrors attached to his eyes.

Challenge 145 s   By the way, why are mirrors used in telescopes, but not in microscopes?

Does light always travel in a straight line? – Refraction

Usually light moves in straight lines. A laser in a misty night shows this most clearly, as illustrated in Figure 76. But any laser pointer in the mist is equally fascinating. Indeed, we use light to *define* 'straightness.' However, there are a number of exceptions that every
Ref. 84   expert on motion should know.
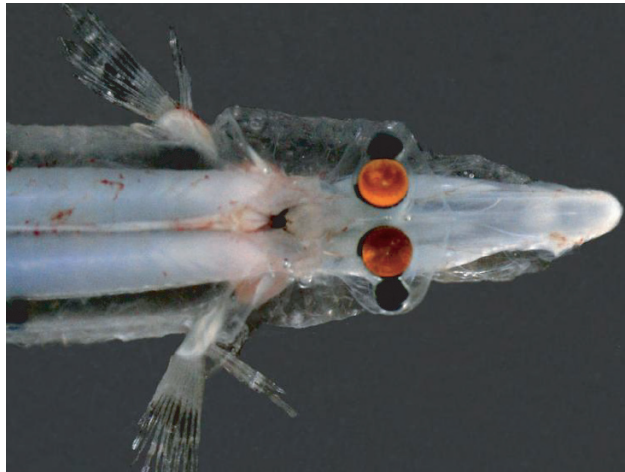
FIGURE 75 The spookfish *Dolichopteryx longipes* has orange mirrors that help him make sharp images also from the dim light coming upwards from bioluminescent lifeforms below it (© Tamara Frank).

FIGURE 76 A sodium frequency laser beam is used as laser guide star to provide a signal for adaptive optics in large telescopes. The laser illuminates a layer of sodium found in the atmosphere at around 90 km of altitude, thus providing an artificial star. The images of the real stars are blurred due to the long exposure time of 3 min (photo Paul Hirst).

In diluted sugar syrup, light beams curve, as shown in Figure 77. The reason is that in such an experiment, the sugar concentration changes with depth. Are you able to explain the syrup effect?

More detailed observation show that a light beam is bent at every material change it encounters on its path. This effect, called *refraction*, is quite common. Refraction changes the appearance of the shape of our feet when we are in the bath tub; refraction also makes aquaria seem less deep than they actually are. Refraction is a consequence of the change of the phase velocity of light from material to material, as shown by Figure 78. The figure also implies that the ratio of the sine of the angles is given by the wave speed ratio, the *refractive index*, usually abbreviated n.

In many fluids and solids, light signals move more slowly than in vacuum; also the (different) phase and group velocities of light inside materials are regularly *lower* than c,
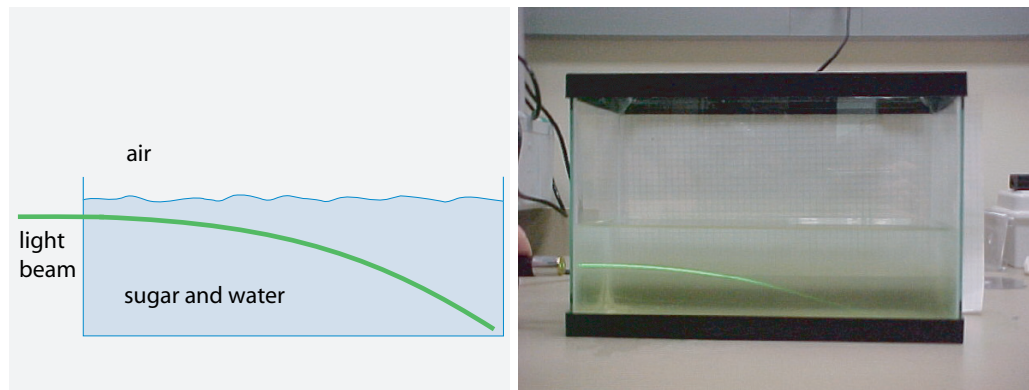
FIGURE 77 Diluted sugar syrup bends light (photo © Jennifer Nierer).
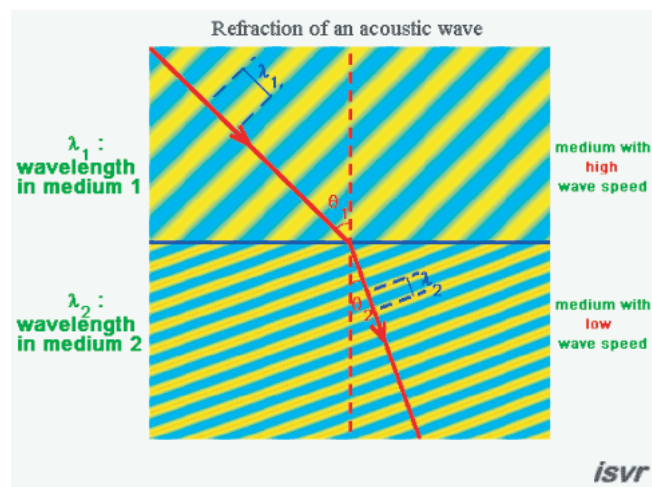
FIGURE 78 A visualisation of refraction (QuickTime film © ISVR, University of Southhampton).

the light speed in vacuum. We discussed the difference between these speeds above. For such 'normal' materials, the refractive index n, the ratio of c to the phase velocity inside the material, is larger than 1. (The difference between c and the phase, group and signal speeds in air are usually neglgigible, so that vacuum and air can usually be interchanged in these discussions.) The refractive index is a material property. For example, the value for visible light in water is about 1.3, for glasses it is around 1.5, and for diamond 2.4. The large value for diamond is one of the reasons for their sparkle. The refractive index also depends on wavelength; prisms make use of this dependence to split white or other light into its constituent colours. Also diamond works as a prism, and this is a second reason for their sparkle.

Many materials however, have refractive indices that are higher than 1, and thus phase velocities larger than c. For example, gold has a refractive index of around 0.2 for visible light, and thus a phase velocity of around $5c$ for such waves. In fact, almost all materials have refractive indices below 1 for *some* wave frequencies, including table salt.

In short, refraction of light, the change of the direction of light motion, is due to dif-
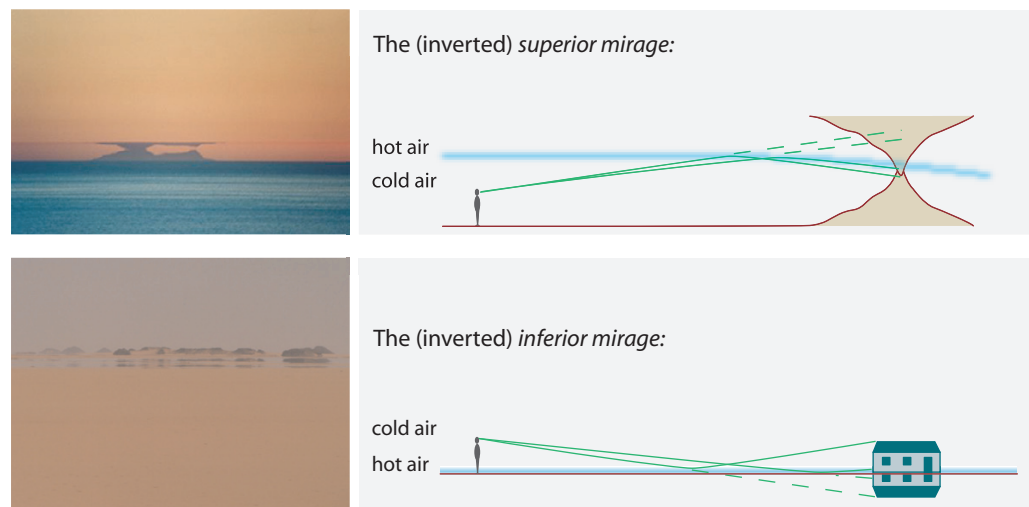
FIGURE 79 The basis of mirages is an effective reflection due to refraction in a hot air layer; it can lead to spectacular effects, such as the inverted superior mirage (top left and right) and the inferior image (bottom left and right) (photographs © Thomas Hogan and Andy Barson).

ferent phase velocities of light in different materials. *Material changes bend light paths.* It is extremely rare to have different materials with the same refractive index.

Also gases lead to refraction. In particular, the refractive index in gases depends on temperature. Refraction in gases leads to the most famous effect of light curvature: the *mirage.* Figure 79 shows photographs of a *superior* mirage, which relies an inversion or boundary layer in the atmosphere above the object and the observer, and a *inferior* mirage, due to a hot layer of air just above the ground. Inferior mirages are also regularly seen on hot highways. All mirage types are due to refraction; their detailed appearance depends on the given temperature profile in the air, and the relative heights of the observer, the inversion layer and the observed object.

Above all, refraction is used in the design of *lenses.* With glass one can produce precisely curved surfaces, allowing light to be *focused.* All focusing devices, such as lenses, can be used to produce images. The two main types of lenses, with their focal points and the images they produce, are shown in Figure 80; they are called *converging lenses* and *divergent lenses.* When an object is more distant from a single converging lens than its focus, the lens produces a *real* image, i.e., an image that can be projected onto a screen. In all other cases single converging or diverging lenses produce so-called *virtual images*: such images can be seen with the eye but not be projected onto a screen. For example, when an object is put between a converging lens and its focus, the lens works as a *magnifying glass.* Figure 80 also allows one to deduce the *thin lens formula* that connects the lengths d_o, d_i and f. What is it?

Even though glasses and lenses have been known since antiquity, the Middle Ages had to pass by before two lenses were combined to make more elaborate optical instruments. The various effects that can be observed with one or two lenses are shown in Figure 81. The *telescope* was invented in, or just before, 1608 by the German–Dutch lens grinder Johannes Lipperhey (c. 1570–1619), who made a fortune by selling it to the Dutch mili-

Ref. 86

Challenge 147 s

Ref. 87
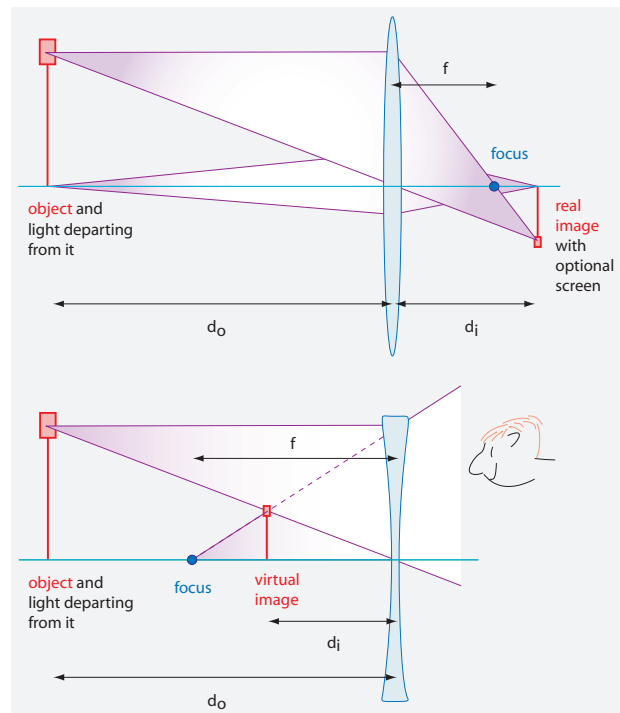
FIGURE 80 A real image produced by a converging lens (if used in the way shown) and the virtual image produced by a diverging lens.

tary. When Galileo heard about the discovery, he quickly took it over and improved it. Already in 1609, Galileo performed the first astronomical observations; they made him world-famous. The *Dutch telescope* design has a short tube yielding a bright and upright image, and its magnification is the ratio of the focal distances of the two lenses. It is still used today in opera glasses. Over the years, many other ways of building telescopes have been developed; in particular, many modern high-performance telescopes use mirrors. Since mirrors are cheaper and easier to fabricate for high-precision imaging, most large telescopes have a mirror instead of the first lens.

By the way, telescopes also exist in nature. Many spiders have two types of eyes. The large ones, made to see distant objects, have two lenses arranged in the same way as in the telescope.

Another way to combine two lenses leads to the *microscope*. Can you explain to a non-physicist how a microscope works? Werner Heisenberg almost failed his Ph.D. exam because he could not. The problem is not difficult, though. Indeed, the inventor of the microscope was an autodidact of the seventeenth century: the Dutch technician Antoni van Leeuwenhoek (1632–1723) made a living by selling over five hundred of his microscopes to his contemporaries. (This is a somewhat nasty remark: Van Leeuwenhoek only used one lens, not two, as in the modern microscope.)

No ray tracing diagram, be it that of a simple lens, of a telescope or of a microscope, is really complete if the eye, with its lens and retina, is missing. Can you add it and convince yourself that these devices really work?

Refraction is usually colour-dependent. For that reason (and also in order to com-
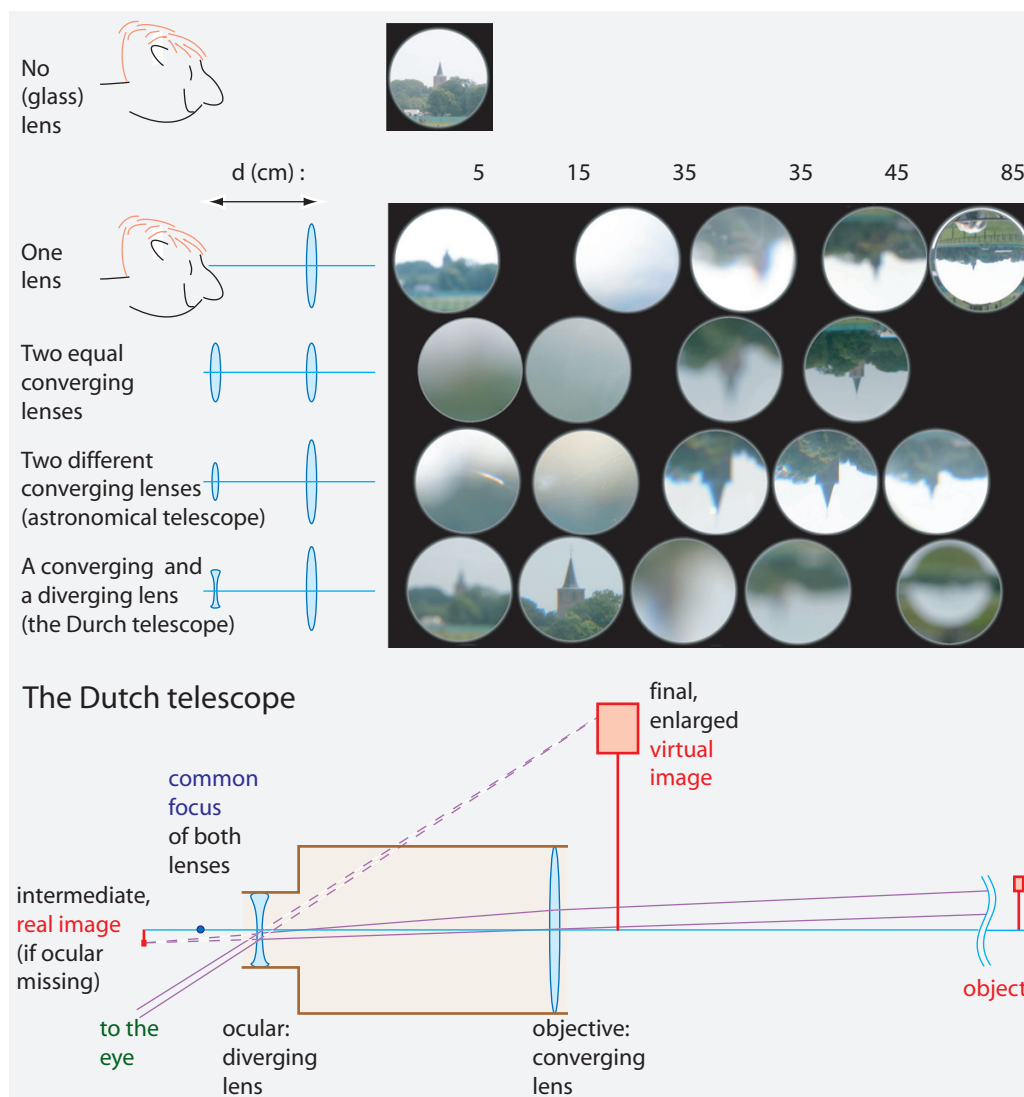
FIGURE 81 Lens refraction as the basis of the telescope: above, the experiments with lenses that lead to the development of the telescope: the object to watch compared with the images produced by a single converging lens, by two equal converging lenses, by two different converging lenses in the astronomical telescope, and by a diverging and a converging lens in the Dutch telescope, at various distances from the eye; below, the explanation of the Dutch telescope (photographs © Eric Kirchner).

pensate the other lens imaging errors called *aberrations*) microscopes or photographic cameras have *several* lenses, made of different types of glass. They compensate the colour effects, which otherwise yield coloured image borders. The colour dependence of refraction in water droplets is also the basis of the rainbow, as shown below, and refraction in ice crystals in the atmosphere is at the basis of the halos and the many other light patterns often seen around the Sun and the Moon.

Also the human eye has colour-dependent refraction. (It is well known that for the working of the eye, the curvature of the cornea is more important than the refractive

FIGURE 82
The glory
produced by
the droplets
in a cloud.

Eye lens dispersion
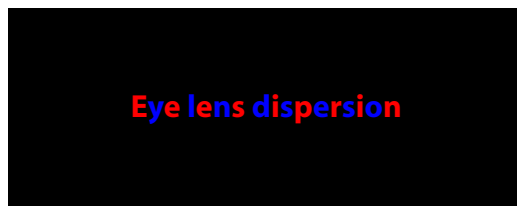
FIGURE 83 Watching this graphic at higher
magnification shows the dispersion of the
human eye: the letters float at different depths.

power of the lens, because the lens is embedded in a medium with nearly the same index
of refraction, thus limiting the effects of refraction.) The colour-dependent refraction is
not corrected in the eye, but in the brain. Indeed, the dispersion of the eye lens can be
noticed if this correction is made impossible, for example when red or blue letters are
printed on a black background, as shown in Figure 83. One gets the impression that the
red letters float in front of the blue letters. Can you show how dispersion leads to the
Challenge 151 s floating effect?

Bending light with tubes – fibre optics

Another way to bend light, also based on refraction, is used by many animals and techni-
cal devices: optical fibres. Optical fibres are based on total internal relection; an overview
of their uses is given in Figure 84.

In nature, optical fibres appear in at least three systems. In insect eyes, such as the
eyes of the house fly or the eye of a honey bee, the light for each pixel is transported
along a structure that works as a conical optical fibre. In certain sea animals, such as the
glass sponge *Euplectella aspergillum* and a number of other sponges, actual silica fibres
are used to provide structural stability and to transport light signals to photodetectors.
Finally, all vertebrate eyes, including the human eye, contain a large number of optical
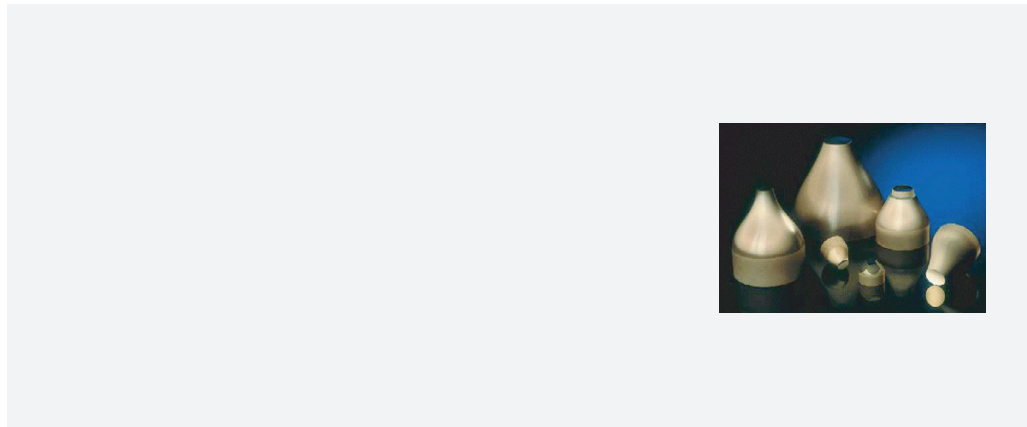
FIGURE 84 Optical fibres: their working principle and their application in the vertebrate eye, in marine
animals, in insect eyes, in illumination, in tapers for image size change, and in telecommunications
(© Schott).
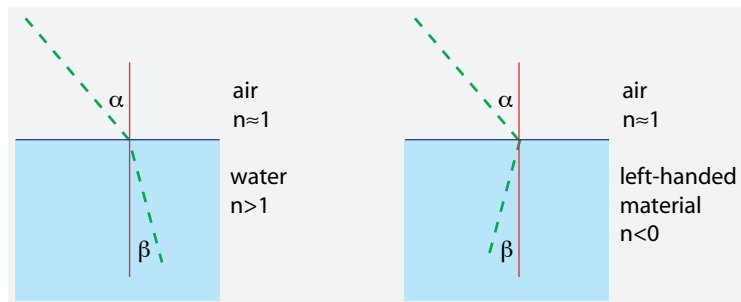
FIGURE 85 Positive and
negative indices of
refraction n.

fibres above the retina, to avoid the image problems that might be caused by the blood
vessels, which lie *above* the retina in all vertebrate eyes. By the way, the frequently heard
claim that the white hair of polar bears works as optical fibres for UV light is *false*.

In technical applications, optical fibres are essential for the working of the telephone
network and the internet, for signal distribution inside aeroplanes and cars, for the trans-
port of laser light for medical uses, for high-power lasers and in many other settings.
Hollow glass fibers are successfully used for the imaging of X-rays.

200 YEARS TOO LATE – NEGATIVE REFRACTION INDICES

In 1968 the Soviet physicist Victor Veselago made a strange prediction: the index of re-
fraction could have *negative* values without invalidating any known law of physics. A
negative index means that a beam is refracted to the same side of the vertical, as shown
in Figure 85. As a result, concave lenses made of such materials focus and convex lenses
disperse parallel beams, in contrast to usual lens materials.

In 1996, John Pendry and his group proposed ways of realizing such materials. In
2000, a first experimental confirmation for microwave refraction was published, but it
met with strong disbelief. In 2002 the debate was in full swing. It was argued that neg-
ative refraction indices imply speeds greater than that of light and are only possible for

Ref. 90
Ref. 91

Ref. 92
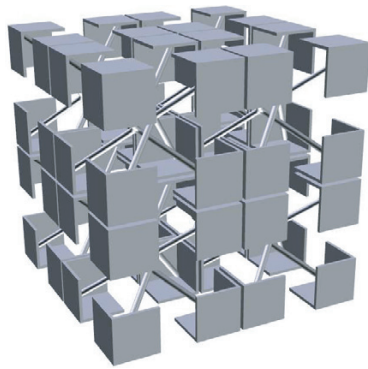
FIGURE 86 An example of an isotropic metamaterial (M. Zedler et al., © 2007 IEEE).

either phase velocity or group velocity, but not for the energy or true signal velocity. The conceptual problems would arise only because in some physical systems the refraction angle for phase motion and for energy motion differ.

In the meantime, the debate is over. Negative indices of refraction have indeed been frequently observed; the corresponding systems are being extensively explored all over the world. Systems with negative index of refraction do exist. The materials showing this property are called *left-handed*. The reason is that the vectors of the electric field, the magnetic field and the wave vector form a left-handed triplet, in contrast to vacuum and most usual materials, where the triplet is right-handed. Such materials consistently have *negative* magnetic permeability and *negative* dielectric coefficient (permittivity).

Apart from the unusual refraction properties, left-handed materials have negative phase velocities, i.e., a phase velocity opposed to the energy velocity; they also show reversed Doppler effect and yield obtuse angles in the Vavilov–Çerenkov effect (emitting Vavilov–Çerenkov radiation in the backward instead of in the forward direction).

But, most intriguing, negative refraction materials are predicted to allow the construction of lenses that are completely flat. In addition, in the year 2000, John Pendry gained the attention of the whole physics community world-wide by predicting that lenses made with such materials, in particular for $n = -1$, would be *perfect*, thus beating the usual diffraction limit. This would happen because such a lens also images the evanescent parts of the waves, by amplifying them accordingly. First experiments seem to confirm the prediction. Exploration of the topic is still in full swing.

It should be mentioned that one type of negative refraction systems have been known since a long time: diffraction gratings. The only difference is that negative index materials try to work as gratings in all spatial directions.

METAMATERIALS

The simplest realization of left-handed systems are metamaterials. *Metamaterials* are engineered metal-insulator structures that are left-handed for a certain wavelength range, usually in the radio or microwave domain.

Currently, there are two basic approaches to realize metamaterials. The first is to build a metamaterial from a large array of resonant inductor-capacitor (LC-) circuits. The second approach is to build them from resonant transmission lines. The latter approach has

Ref. 93

Ref. 94

Ref. 95

Ref. 94

Ref. 96

Ref. 97

lower losses and a wider spectral range. An example is shown in Figure 86.

Most metamaterials are conceived for waves in the microwave range. Crystals that act as left-handed materials in the optical range, though for certain frequencies only, and only in one dimension, have been identified recently. Research on the topic is presently in full swing. Industrial applications of left handed (meta-)materials are expected for antenna design; for example, the antenna dipole could be located just above a metamaterial, allowing flat directional antennas. The most unrealistic people are those who claim that *invisibility cloaks* can be realized with metamaterials. While this is a good marketing slogan to attract funding, it is not realistic, due to inevitable signal losses in the materials, dispersion, finite cell size etc. So far, all aeroplanes that were claimed to be invisible for certain (radar) frequencies have turned out to be visible to radar. But sources of military funding are known to have only a distant relation to reality.

LIGHT AROUND CORNERS – DIFFRACTION

Light goes around corners. This effect was called *diffraction* by Francesco Grimaldi, in his text *Physico-mathesis de lumine*, published in 1665. Grimaldi studied shadows very carefully. He found out what everybody now learns in secondary school: light goes around corners in the same way that sound does, and light diffraction is due to the wave nature of light. (Newton got interested in optics after he read Grimaldi; he then wrongly dismissed his conclusions.)

Because of diffraction, it is *impossible* to produce strictly parallel light beams. For example, every laser beam diverges by a certain minimum amount, called the *diffraction limit*. Maybe you know that the world's most expensive Cat's-eyes are on the Moon, where they have been deposited by the Lunokhod and the Apollo missions. Can you determine how wide a laser beam with minimum divergence has become when it arrives at the Moon and returns back to Earth, assuming that it was 1 m wide when it left Earth? How wide would it be when it came back if it had been 1 mm wide at the start? In short, diffraction and the impossibility of parallel beams confirms that light is a wave.

Diffraction implies that there are no perfectly sharp images: there exists a *limit on resolution*. This is true for every optical instrument, including the eye. The resolution of the eye is between one and two minutes of arc, i.e., between 0.3 and 0.6 mrad. The limit is partly due to the finite size of the pupil. (That is why squinting helps to see more sharply.) In practice, the resolution of the eye is often limited by chromatic aberrations and shape imperfections of the cornea and lens. (Can you check the numbers and their interpretation by calculation? Is it true that the number of cones in the eye is tuned exactly to its resolution?) Therefore, for example, there is a maximum distance at which humans can distinguish the two headlights of a car. Can you estimate it?

Resolution limits also make it impossible to see the Great Wall in northern China from the Moon, contrary to what is often claimed. In the few parts that are not yet in ruins, the wall is about 6 metres wide, and even if it casts a wide shadow during the morning or the evening, the angle it subtends is way below a second of arc, so that it is completely invisible to the human eye. In fact, three different cosmonauts who travelled to the Moon performed careful searches and confirmed that the claim is absurd. The story is one of the most tenacious urban legends. (Is it possible to see the Wall from the space shuttle?) The largest human-made objects are the polders of reclaimed land in the Netherlands; they
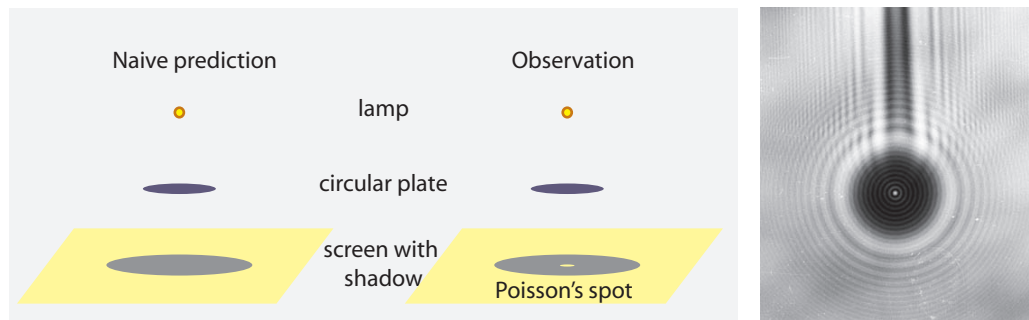
FIGURE 87 Shadows show that light is a wave: the naïve expectation (left), neglecting the wave idea, and the actual observation (middle and right) of the shadow of a circular object (photo © Christopher Jones).

are visible from outer space. So are most large cities as well as the highways in Belgium at night; their bright illumination makes them stand out clearly from the dark side of the Earth.

Diffraction has the consequence that behind a small disc illuminated along its axis, the centre of the shadow shows, against all expectations, a bright spot, as shown in Figure 87. This 'hole' in the shadow was predicted in 1819 by Denis Poisson (1781–1840) in order to show to what absurd consequences the wave theory of light would lead. He had just read the mathematical description of diffraction developed by Augustin Fresnel* on the basis of the wave description of light. But shortly afterwards, François Arago (1786–1853) actually observed Poisson's spot, converting Poisson, making Fresnel famous and starting the general acceptance of the wave properties of light.

In summary, diffraction is not often used to form images; above all, diffraction determines the quality of every image.

BEATING THE DIFFRACTION LIMIT

In all imaging methods, the race is for images with the highest resolution possible. In particular, the techniques of producing images with resolutions *less* than the wavelength of light have made great progress in recent years.

Nowadays, extraordinary images can be produced with modified commercial light microscopes. The conventional diffraction limit for microscopes is

$$d \geqslant \frac{\lambda}{2n \, \sin \alpha} \, , \tag{77}$$

where λ is the wavelength, n the index of refraction and α is the angle of observation.

A recent technique, called *stimulated emission depletion microscopy*, allows spot sizes of molecular size. The new technique is a special type of fluorescence microscopy and

---

* Augustin Jean Fresnel (1788–1827), engineer and part time physicist; he published in 1818 his great paper on wave theory for which he got the prize of the French Academy of Sciences in 1819. (The 's' in his name is silent.) To improve his finances, he worked in the commission responsible for lighthouses, for which he developed the well-known Fresnel lens. He died prematurely, partly of exhaustion due to overwork.
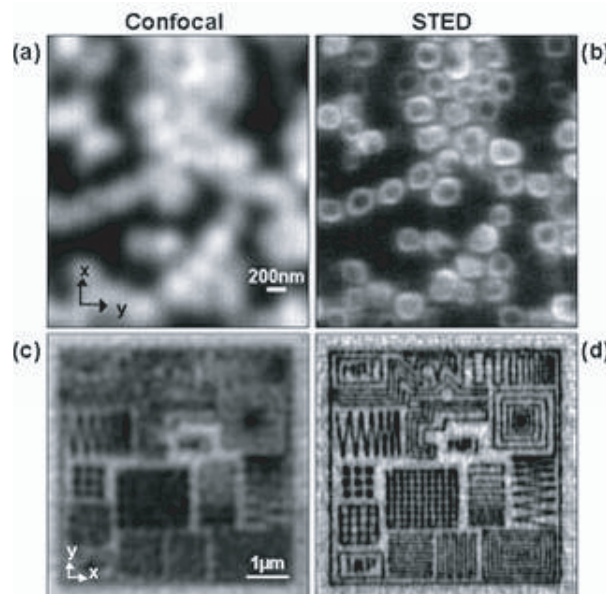
FIGURE 88 Sub-wavelength optical microscopy using stimulated emission depletion (right) compared to conventional confocal microscopy (left) (© MPI für biophysikalische Chemie/Stefan Hell).

preliminary drawing

FIGURE 89 In certain materials, light beams can spiral around each other.
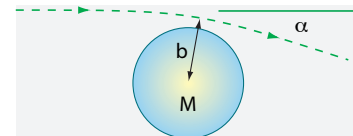
FIGURE 90 Masses bend light.

was developed by Stefan Hell. It modifies the diffraction limit to

$$ d \geqslant \frac{\lambda}{2n \ \sin \alpha \ \sqrt{I/I_{\text{sat}}}} \ , \tag{78} $$

so that a properly chosen saturation intensity allows one to reduce the diffraction limit to arbitrary low values. So far, light microscopy with a resolution of 16 nm has been performed, as shown in Figure 88. This and similar techniques are becoming common-place in medicine, biology and material science.

Ref. 101

MORE WAYS TO BEND LIGHT

Additional electromagnetic fields usually do not influence light directly, since light has no charge and since Maxwell's equations are linear. But in some materials the equations are non-linear, and the story changes. For example, in certain photorefractive materials, two nearby light beams can even *twist* around each other, as was shown by Segev and coworkers in 1997. This is illustrated in Figure 89.

Ref. 102

A further way to bend light is gravity, as discussed already in the chapters on universal gravity and on general relativity. Since the effect is weak, it is only of importance in as-
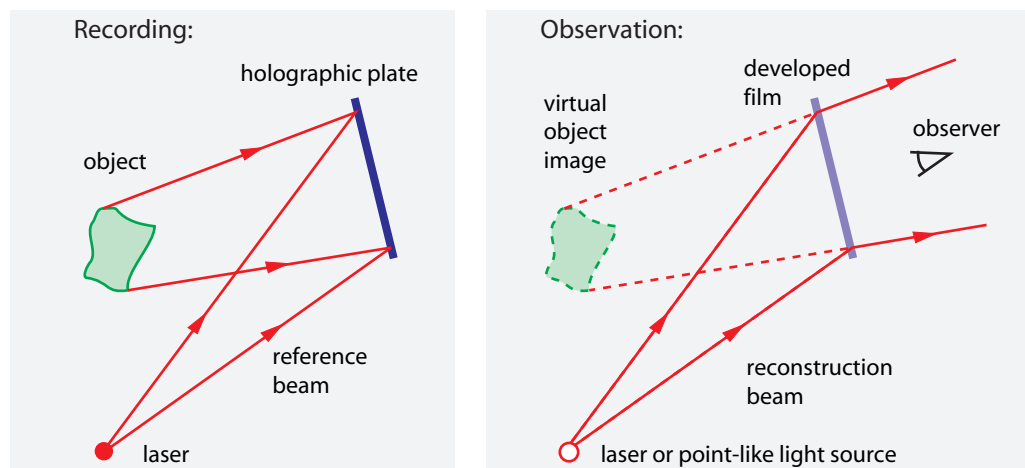
FIGURE 91 The recording (left) and the observation (right) of a hologram.

tronomy. Gravitational lensing is used in various projects to measure the size, mass and distance of galaxies and galaxy groups. The usually negligible effect of gravity between two light beams was also discussed already.

A light ray that is scattered also changes direction, though it is debatable whether it is appropriate to call this process an example of bending of light. In any case, scattering is important: without it, we would not see almost anything around us. After all, everyday seeing is detection of scattered light.

In summary, light travels straight only if it travels *far from matter*. In everyday life, 'far' simply means more than a few millimetres, because all gravitational and electromagnetic effects are negligible at these distances, mainly due to light's truly supersonic speed. However, as we have seen, in some cases larger distances from matter are necessary to ensure undisturbed motion of light.

HOW DOES ONE MAKE HOLOGRAMS AND OTHER THREE-DIMENSIONAL IMAGES?

Our sense of sight gives us the impression of depth mainly due to three effects. First, the two eyes see different images. Second, the images formed in the eyes are position dependent. Third, our eye needs to focus differently for different distances.

A usual photograph does not capture any three-dimensional effect. A usual photograph corresponds to the picture taken by one eye, from one particular spot and at one particular focus. In fact, all photographic cameras are essentially copies of a single static eye.

Any system wanting to produce the perception of depth must include at least one of the three effects just mentioned. In all systems so far, the third and weakest effect, varying focus with distance, is never used, as it is too weak. Stereo photography and virtual reality systems extensively use the first effect by sending two different images to the eyes. Also certain post cards and computer screens are covered by thin cylindrical lenses that allow one to send two different images to the two eyes, thus generating the same impression of depth.

But obviously the most spectacular effect is obtained whenever position dependent images can be created. Some virtual reality systems mimic this effect using a sensor attached to the head, and creating computer–generated images that depend on this position. However, such systems are not able to reproduce actual situations and thus pale when compared with the impression produced by holograms.

Holograms reproduce all that is seen from any point of a region of space. A *hologram* is thus a stored set of position dependent pictures of an object. It is produced by storing amplitude *and phase* of the light emitted by an object. To achieve this, the object is illuminated by a *coherent* light source, such as a laser, and the interference pattern is stored. Illuminating the developed photographic film by a coherent light source then allows one to see a full three-dimensional image. In particular, due to the reproduction of the situation, the image appears to float in free space. Holograms were developed in 1947 by the Hungarian physicist Dennis Gabor (1900–1979), who received the 1971 Nobel Prize for physics for this work.

Holograms can be made to work in reflection or transmission. The simplest holograms use only one wavelength. Most coloured holograms are rainbow holograms, showing false colours that are unrelated to the original objects. Real colour holograms, made with three different lasers, are rare but possible.

Is it possible to make *moving* holograms? Yes; however, the technical set-ups are still extremely expensive. So far, they exist only in a few laboratories and cost millions of euro. By the way, can you describe how you would distinguish a moving hologram from a real body, if you ever came across one, without touching it?

Tomography

A spectacular type of imaging has become possible only after high-speed computers became cheap: *tomography*. In tomography, a radiation source rotates around the object to be imaged; the radiation that is scattered and/or trasmitted is detected, and with sophisticated computer programming, a cross section of the object is reconstructed. Three-dimensional reconstructions are also possible. Tomography can be performed with any type of radiation that can be emitted in precisely defined beams, such as X-rays, light of various frequencies, and even sound. X-ray tomography is a standard method in health care; visible light tomography, which has no side effects, is being developed for breast tumor detection.

In many types of tomography, the resolution achieved is becoming spectacular. An example for modern high-resolution X-ray tomography of really small objects is shown in Figure 92. Also magnetic resonance imaging, widely used in health care to image the interiour of the human body, is a type of tomography; it will be presented later on in our journey. Various types of tomography also allow the production of films.

The most exotic method is *myon tomography*, an imaging method that uses the myons in cosmic rays to detect heavy metals in boxes, luggage and trucks. This method is particularly interesting for searching for hidden heavy metals, such as plutonium, which scatter myons much more strongly than other materials such as iron.
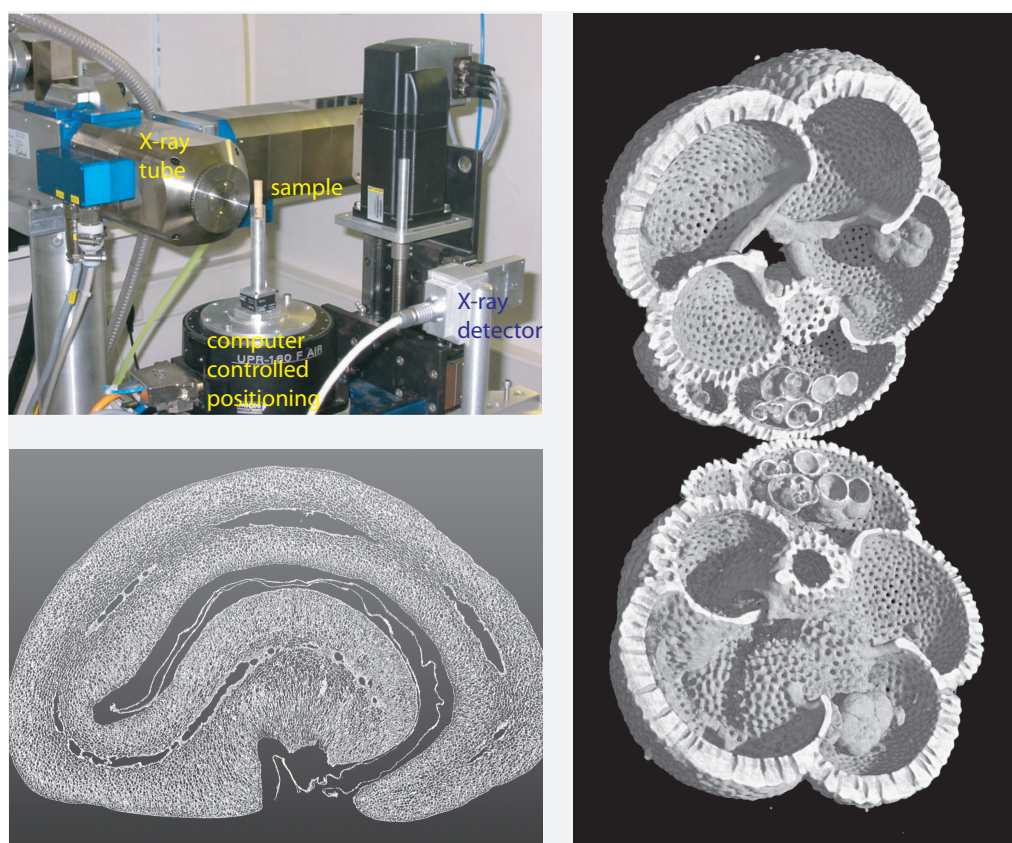
FIGURE 92 A set-up for high-resolution x-ray tomography, and two examples of images produced with it: a cross-section of a coffee bean (lower left) with a size of 8 mm, and a three-dimensional reconstruction of the exoskeleton of a foraminiferan, with a diameter of only 0.5 mm (© Manuel Dierick).

THE EYE AND THE BRAIN: OBSERVING IMAGES

DO WE SEE WHAT EXISTS?

Sometimes we see *less* than there is. Close your left eye, look at the white spot in Figure 93, bring the page slowly towards your eye, and pay attention to the middle lines. At a distance of about 15 to 20 cm the middle line will seem uninterrupted. Why?

Challenge 158 s

On the other hand, sometimes we see *more* than there is, as Figures 94 and 95 show. The first shows that parallel lines can look skewed, and the second show a so-called *Hermann lattice*, named after its discoverer.* The Hermann lattice of Figure 95, discovered by Elke Lingelbach in 1995, is especially striking. Variations of these lattices are now used to understand the mechanisms at the basis of human vision. For example, they can be used to determine how many light sensitive cells in the retina are united to one signal pathway towards the brain. The illusions are angle dependent because this number is also angle dependent.

Ref. 103

---

* Ludimar Hermann (1838–1914), Swiss physiologist. The lattices are often falsely called 'Hering lattices' after the man who made Hermann's discovery famous.

FIGURE 93 A limitation of the eye (see text).
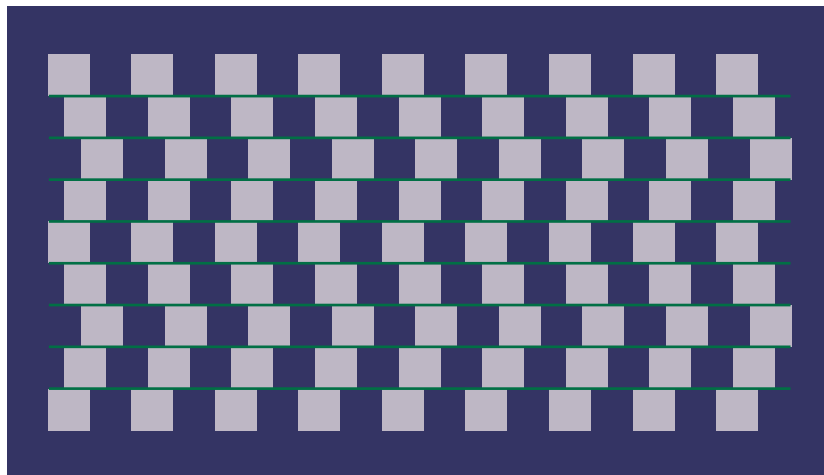
FIGURE 94 What is the angle between the thin lines between the squares?

Our eyes also 'see' things *differently:* the retina sees an *inverted* image of the world. There is a simple method to show this, due to Helmholtz.* You need only a needle and a piece of paper, e.g. this page of text. Use the needle to make two holes inside the two letters 'oo'. Then keep the page as close to your eye as possible, look through the two holes towards the wall, keeping the needle vertical, a few centimetres behind the paper. You will see two images of the needle. If you now cover the *left* hole with your finger, the *right* needle will disappear, and vice versa. This shows that the image inside the eye, on

---

* See HERMANN VON HELMHOLTZ, *Handbuch der physiologischen Optik*, 1867. This famous classic is available in English as *Handbook of Physiological Optics*, Dover, 1962. The Prussian physician, physicist and science politician born as Hermann Helmholtz (b. 1821 Potsdam, d. 1894 Charlottenburg) was famous for his works on optics, acoustics, electrodynamics, thermodynamics, epistemology and geometry. He founded several physics institutions across Germany. He was one of the first to propagate the idea of conservation of energy. His other important book, *Die Lehre von den Tonempfindungen*, published in 1863, describes the basis of acoustics and, like the Handbook, is still worth reading.
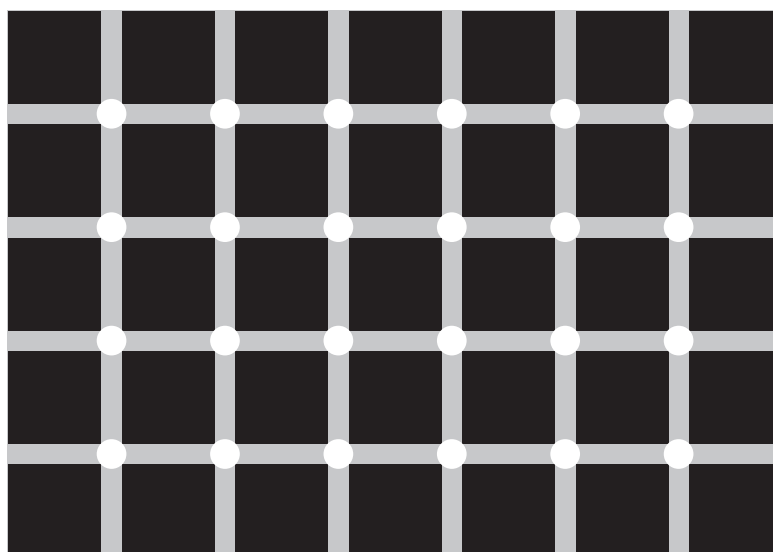
FIGURE 95 The Lingelbach lattice: do you see white, grey, or black dots at the crossings?

FIGURE 96 An example of an infrared photograph, slightly mixed with a colour image (© Serge Augustin).

the retina, is inverted. Are you able to complete the proof?

Two other experiments can show the same result. If you push very lightly on the *inside* of your eye (careful!), you will see a dark spot appear on the *outside* of your vision field. And if you stand in a dark room and ask a friend to look at a burning candle, explore his eye: you will see three reflections: two upright ones, reflected from the cornea and from the lens, and a dim third one, *upside-down*, reflected from the retina.

Another reason that we do not see the complete image of nature is that our eyes have a limited sensitivity. This sensitivity peaks around 560 nm; outside the red and the violet, the eye does not detect radiation. We thus see only part of nature. For example, infrared photographs of nature, such as the one shown in Figure 96, are interesting because they

FIGURE 97 How the appearance of a sunflower changes with wavelength: how it looks to the human eye, how it might look to a bird, and how it looks in the ultraviolet (© Andrew Davidhazy).

show us something different from what we see usually. The same happens to ultraviolet
Ref. 104 photographs, as shown in Figure 97.
      Every expert of motion should also know that the sensitivity of the eye does *not* cor-
Ref. 105 respond to the brightest part of sunlight. This myth has been spread around the world
by the numerous textbooks that have copied from each other. Depending on whether
frequency or wavelength or wavelength logarithm is used, the solar spectrum peaks at
500 nm, 880 nm or 720 nm. The human eye's spectral sensitivity, like the completely dif-
ferent sensitivity of birds or frogs, is due to the chemicals used for detection. In short,
the human eye can only be understood by a careful analysis of its particular evolution-
ary history.
      An urban legend, spread by many medical doctors and midwives to this day, pre-
tends that newborn babies see everything upside down. Can you explain why this idea
Challenge 160 s is wrong?
      In summary, we have to be careful when maintaining that seeing means observing.
Examples such as these lead to ask whether there are other limitations of our senses
which are less evident. And our walk will indeed uncover several of them.

How can one make pictures of the inside of the eye?

The most beautiful pictures so far of a *living* human retina, such as that of Figure 98, were
made by the group of David Williams and Austin Roorda at the University at Rochester in
Ref. 106 New York. They used adaptive optics, a technique that changes the shape of the imaging
lens in order to compensate for the shape variations of the lens in the human eye.*
      The eyes see colour by averaging the intensity arriving at the red, blue and green sensi-
Page 104 tive cones. This explains the possibility, mentioned above, of getting the same impression
of colour, e.g. yellow, either by a pure yellow laser beam, or by a suitable mixture of red
and green light.
      But if the light is focused on to one cone only, the eye makes mistakes. If, using this

---

* Nature uses another trick to get maximum resolution: the eye continuously performs small movements, called *micronystagmus*. The eye continuously oscillates around the direction of vision with around 40 to 50 Hz. In addition, this motion is also used to allow the cells in the retina to recharge.
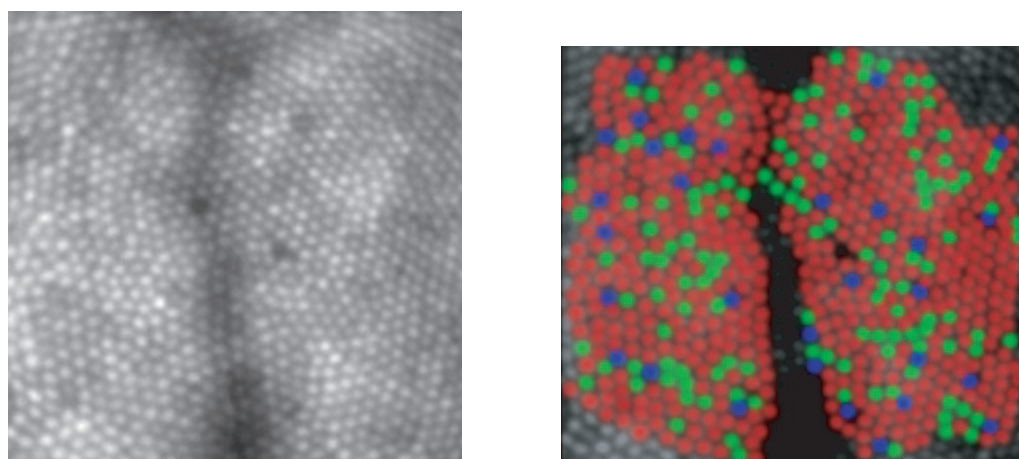
FIGURE 98 A high quality photograph of a live human retina, including a measured (false colour) indication of the sensitivity of each cone cell (© Austin Roorda).
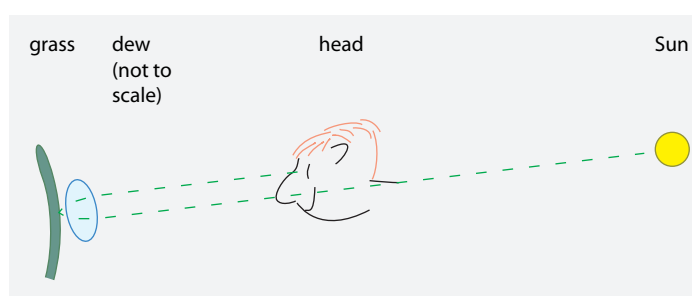
FIGURE 99 The path of light for the dew on grass that is responsible for the aureole.

adaptive optics, a red laser beam is focused such that it hits a green cone only, a strange thing happens: even though the light is *red*, the eye sees a *green* colour!

Incidentally, Figure 98 is quite puzzling. In the human eye, the blood vessels are located *in front* of the cones. Why don't they appear in the picture? And why don't they disturb us in everyday life? (The picture does not show the other type of sensitive light cells, the *rods*, because the subject was in daylight; rods come to the front of the retina only in the dark, and then produce black and white pictures.)

Challenge 161 s

Of all the mammals, only primates can see *colours*. Bulls for example, don't; they cannot distinguish red from blue. On the other hand, the best colour seers overall are the birds. They have cone receptors for red, blue, green, UV and, depending on the bird, for up to three more sets of colours. A number of birds (but not many) also have a better eye resolution than humans. Several birds also have a faster temporal resolution: humans see continuous motion when the images follow with 30 to 70 Hz (depending on the image content); some insects can distinguish images up to 300 Hz.

How to prove you're holy

Light reflection and refraction are responsible for many effects. The originally Indian symbol of holiness, now used throughout most of the world, is the *aureole*, also called

halo or *Heiligenschein*, a ring of light surrounding the head. You can easily observe it around your own head. You need only to get up early in the morning and look into the wet grass while turning your back to the Sun. You will see an aureole around your shadow. The effect is due to the morning dew on the grass, which reflects the light back predominantly in the direction of the light source, as shown in Figure 99. The fun part is that if you do this in a group, you will see the aureole around only *your own* head.

Ref. 107

Retroreflective paint works in the same way: it contains tiny glass spheres that play the role of the dew. A large surface of retroreflective paint, a traffic sign for example, can also show your halo if the light source is sufficiently far away. Also the so-called 'glow' of the eyes of a cat at night is due to the same effect; it is visible only if you look at the cat with a light source behind you. By the way, do Cat's-eyes work like a cat's eyes?

Ref. 108

Challenge 162 s

Challenges and fun curiosities about images and the eye

The eye is a wonderful organ. To learn more about it, read the beautiful book Si-mon Ings, *The Eye – A Natural History*, Bloomsbury, 2007.

* *

The Sun is visible to the naked eye only up to a distance of 50 light years. Is this true?

Challenge 163 s

* *

Grass is usually greener on the other side of the fence. Can you give an explanation based on observations for this statement?

Ref. 109

Challenge 164 s

* *

It is said that astronomers have telescopes so powerful that they could see whether some-body would be lighting a match on the Moon. Can this be true?

Challenge 165 s

* *

When two laser beams cross at a small angle, they can form light pulses that seem to move faster than light. Does this contradict special relativity?

Ref. 110

Challenge 166 s

* *

Colour blindness was discovered by the great English scientist John Dalton (1766–1844) – on himself. Can you imagine how he found out? It affects, in all its forms, one in 20 men. In many languages, a man who is colour blind is called *daltonic*. Women are almost never daltonic, as the property is linked to defects on the X chromosome. If you are colour blind, you can check to which type you belong with the help of Figure 101.

Challenge 167 ny

Ref. 111

* *

Artificial colour blindness is induced by certain types of illumination. For example, violet light is used to reduce intravenous drug consumption, because violet light does not allow finding veins under the skin.

Artificial contrast enhancement with illumination is also useful. Pink light is used by beauticians to highlight blemishes, so that the skin can be cleaned as well as possible. In 2007, the police officer Mike Powis in Nottingham discovered that this 'acne light' could
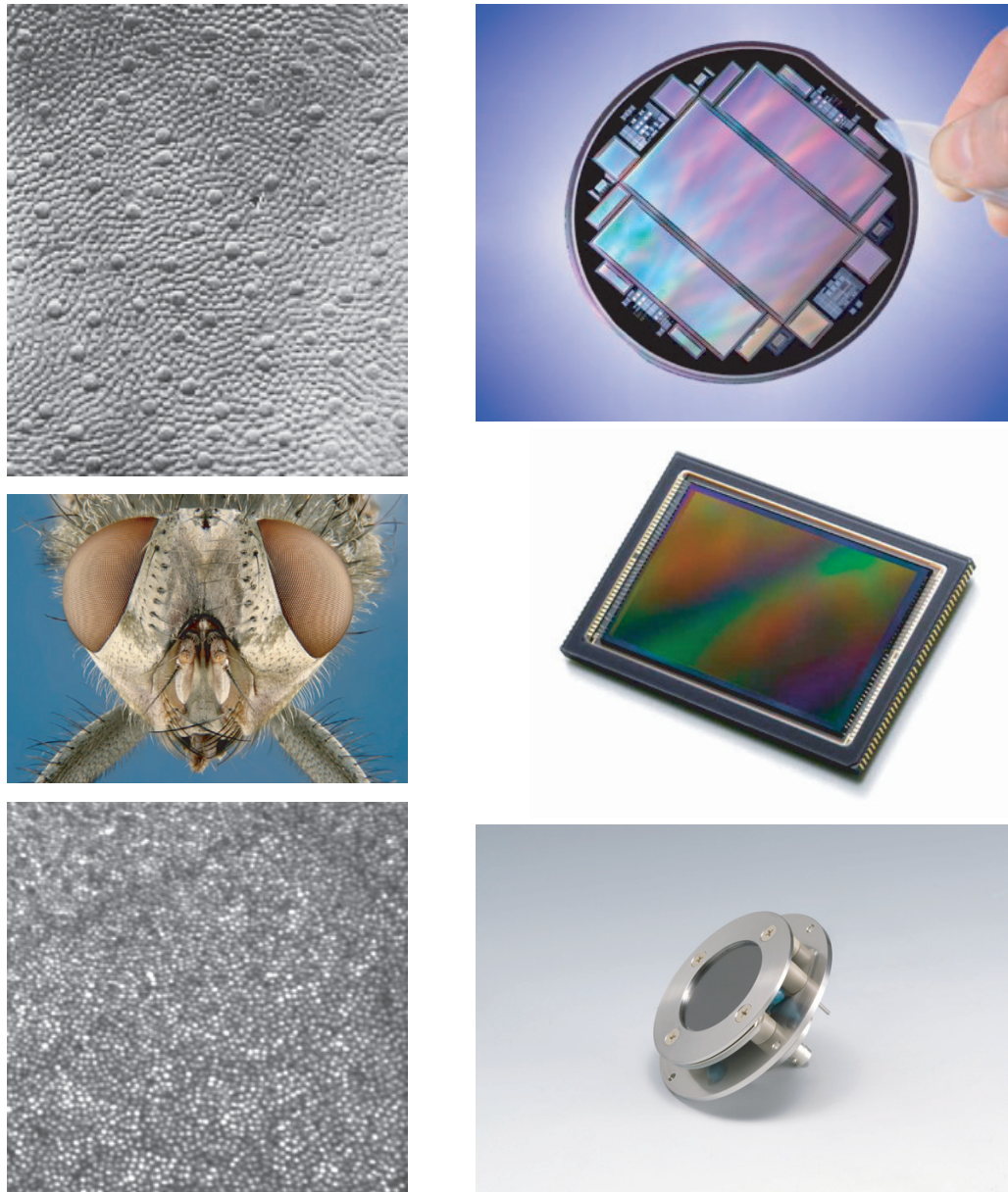
FIGURE 100 A collection of image sensors – thus of pixel systems: A cat's retina, a CCD sensor still on a wafer, the eye of a house fly, a CMOS sensor, a human retina, and a multichannel plate (© Wikimedia, Austin Roorda, Hamamatsu Photonics).

be used to reduce the crime rate; since acne is not fashionable, pink light deters youth from gathering in groups, and thus calms the environment where it is installed.

Yellowish light is used by by supermarkets to increase their sales of fruits and vegetables. In yellow light, tomatoes look redder and salad looks greener. Check by yourself: you will not find a single supermarket without these lights installed over fruits and vegetables.

Challenge 168 e

FIGURE 101 How natural colours (top) change for three types of colour blind: deutan, protan and tritan (© Michael Douma).

* *

Light beams, such as those emitted from lasers, are usually thought of as thin lines. However, light beams can also be *tubes*. Tubular laser beams, i.e., Bessel beams of high order, are used in modern research experiments to guide plasma channels and sparks.

* *

Is it possible to see stars from the bottom of a deep pit or of a well, even during the day, as is often stated?

* *

Imaging is an important part of modern industry. Without laser printers, photocopying machines, CD players, DVD players, microscopes, digital photographic cameras, film and video cameras, lithography machines for integrated circuit production, telescopes, film projectors, our world would look much more boring. And there are not enough experts in the world on this fascinating field.

* *

The simplest imaging system are eye glasses. A child that has no proper glasses misses an important experience: seeing the stars. Such a child will not understand the famous statement by Immanuel Kant: 'Two things fill the mind with ever new and increasing admiration and awe, the more often and persistently thought considers them: the starred sky above me and the moral law inside me.' Always be sure that children can see the stars.

* *

Humans are the *only* primates that have *white* eyes. All apes have *brown* eyes, so that it is impossible to see in which direction the are looking. Apes make extensive use of this impossibility: they often turn their head in one direction, pretending to look somewhere,

FIGURE 102 Ames rooms in Paris and in San Francisco (© Sergio Davini, David Darling).

but turn their eyes in another. In other words, brown eyes are useful for deception. The same effect is achieved in humans by wearing dark sunglasses. So if you see somebody with sunglasses in a situation where there is no sunlight, you know that he or she is behaving like an ape.

* *

Challenge 170 s   How can you measure the power of the Sun with your eyes closed?

* *

Even in a dark, moonless and starless night, a forest is not dark. You can see luminescent mushrooms (of which there are over 70 different species), luminescent moulds, you can see sparks when you take off your pullover or when your friend bites a mint bonbon or Challenge 171 ny   when you unroll a roll of adhesive tape or open a letter. What else can you see?

* *

The number of optical illusions is enormous, and there are many websites devoted to the topic. A pretty illusion, often used in films to transform actors into dwarfs, is the so-called *Ames room*, shown in Figure 102.

* *

Page 140   The eye and the brain sometimes *add* false colours, as we have seen above in the discussion of rods. Also Haidinger's brush is an example of added colours. In contrast, sometimes the brain and the eye make colours *disappear*, as shown in Figure 103. The example is taken from the beautiful collection of visual illusions at www.psy.ritsumei.ac.jp/~akitaoka/color9e.html.

Summary on optics

The art and science of making images is in large part a result of bending light beams in predefined ways. To achieve high quality images, material science, mechatronics, signal processing, and many other engineering domains are required.
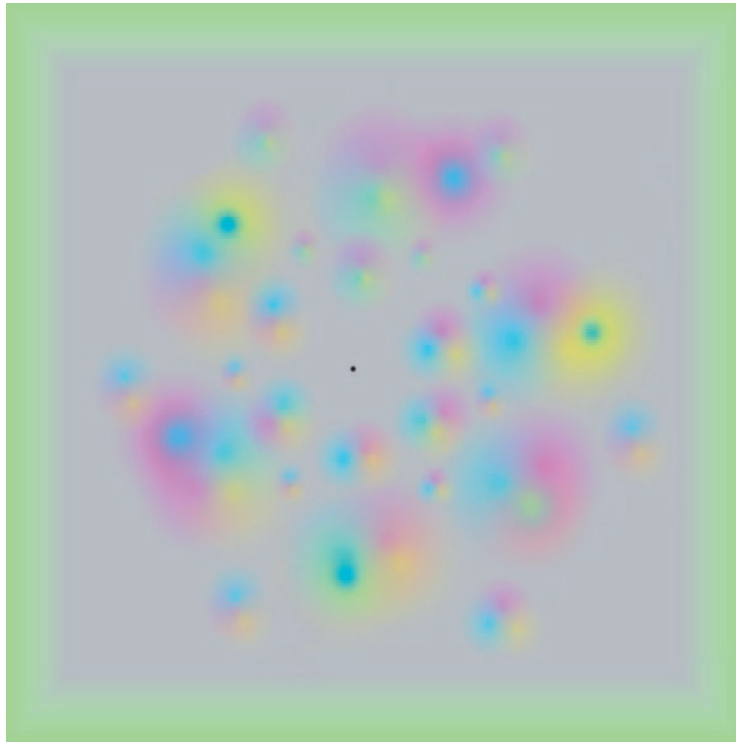
FIGURE 103 Look at the central dot for twenty seconds: the colours will disappear (© Kitaoka Akiyoshi).

CHAPTER 5

ELECTROMAGNETIC EFFECTS
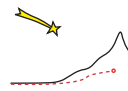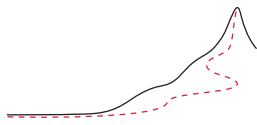
L OOKING carefully, the atmosphere is full of electrical effects. The most impressive, lightning, is now reasonably well understood. However, it took decades and a large number of researchers to discover and put together all the parts of the puzzle. Also below our feet there is something important going on: the hot magma below the continental crust produces the magnetic field of the Earth. Strong magnetic fields can be used for levitation. We explore these topics first and then give an overview about the many effects that fields produce.

Is lightning a discharge? – Electricity in the atmosphere

Inside a thunderstorm cloud, especially inside tall *cumulonimbus* clouds,* charges are separated by collision between the large 'graupel' ice crystals falling due to their weight and the small 'hail' ice crystallites rising due to thermal upwinds. Since the collision takes part in an electric field, charges are separated in a way similar to the mechanism in the Kelvin generator. Discharge takes place when the electric field becomes too high, taking a strange path influenced by ions created in the air by cosmic rays. (There are however, at least ten other competing explanations for charge separation in clouds.) It seems that cosmic rays are at least partly responsible for the zigzag shape of lightning.

Lightning flashes have strange properties. First, they appear at fields around 200 kV/m (at low altitude) instead of the 2 MV/m of normal sparks. Second, lightning emits radio pulses. Third, they emit gamma rays. Russian researchers, from 1992 onwards explained all three effects by a newly discovered discharge mechanism. At length scales of 50 m and more, cosmic rays can trigger the appearance of lightning; the relativistic energy of these rays allows for a discharge mechanism that does not exist for low energy electrons. At relativistic energy, so-called runaway breakdown leads to discharges at much lower fields than usual laboratory sparks. The multiplication of these relativistic electrons also leads to the observed radio and gamma ray emissions.

Ref. 114

Page 19

Ref. 115

Ref. 116

---

* Clouds have Latin names. They were introduced in 1802 by the English explorer Luke Howard (1772–1864), who found that all clouds could be seen as variations of three types, which he called *cirrus*, *cumulus* and *stratus*. He called the combination of all three, the rain cloud, *nimbus* (from the Latin 'big cloud'). Today's internationally agreed system has been slightly adjusted and distinguishes clouds by the height of their lower edge. The clouds starting above a height of 6 km are the cirrus, the cirrocumulus and the cirrostratus; those starting at heights of between 2 and 4 km are the altocumulus, the altostratus and the nimbostratus; clouds starting below a height of 2 km are the stratocumulus, the stratus and the cumulus. The rain or thunder cloud, which crosses all heights, is today called *cumulonimbus*. For beautiful views of clouds, see the www.goes.noaa.gov and www.osei.noaa.gov websites.

Ref. 113

FIGURE 104 Cumulonimbus clouds from ground and from space (NASA).

Incidentally, you have a 75 % chance of survival after being hit by lightning, especially if you are completely wet, as in that case the current will flow outside the skin. Usually, wet people who are hit lose all their clothes, as the evaporating water tears them off. Rapid resuscitation is essential to help somebody to recover after a hit.[*]

As a note, you might know how to measure the distance of a lightning by counting the seconds between the lightning and the thunder and multiplying this by the speed of sound, 330 m/s; it is less well known that one can estimate the *length* of the lightning bolt by measuring the *duration* of the thunder, and multiplying it by the same factor.

In the 1990s more electrical details about thunderstorms became known. Airline pilots and passengers sometime see weak and coloured light emissions spreading from the top of thunderclouds. There are various types of such emissions: blue *jets* and mostly red *sprites* and *elves*, which are somehow due to electric fields between the cloud top and the ionosphere. The details are still under investigation, and the mechanisms are not yet clear.[**]

All these details are part of the electrical circuit around the Earth. This fascinating part of geophysics would lead us too far from the aim of our mountain ascent. But every physicist should know that there is a vertical electric field of between 100 and 300 V/m on a clear day, as discovered already in 1752. (Can you guess why it is not noticeable in everyday life? And why despite its value it cannot be used to extract large amounts of energy?) The field is directed from the ionosphere down towards the ground; in fact the Earth is permanently negatively charged, and in clear weather current flows downwards through the clear atmosphere, trying to *discharge* our planet. The current of about 1 kA is spread over the whole planet; it is possible due to the ions formed by cosmic radiation. (The resistance between the ground and the ionosphere is about 200 Ω, so the total voltage drop is about 200 kV.) At the same time, the Earth is constantly being *charged* by several effects: there is a dynamo effect due to the tides of the atmosphere and there are currents induced by the magnetosphere. But the most important effect is lightning. In

Challenge 172 s

---

[*] If you are ever hit by lightning and survive, go to the hospital! Many people died three days later having failed to do so. A lightning strike often leads to coagulation effects in the blood. These substances block the kidneys, and one can die three days later because of kidney failure. The remedy is to have dialysis treatment.
[**] For images, have a look at the interesting sprite.gi.alaska.edu/html/sprites.htm, www.fma-research.com/spriteres.htm and paesko.ee.psu.edu/Nature websites.
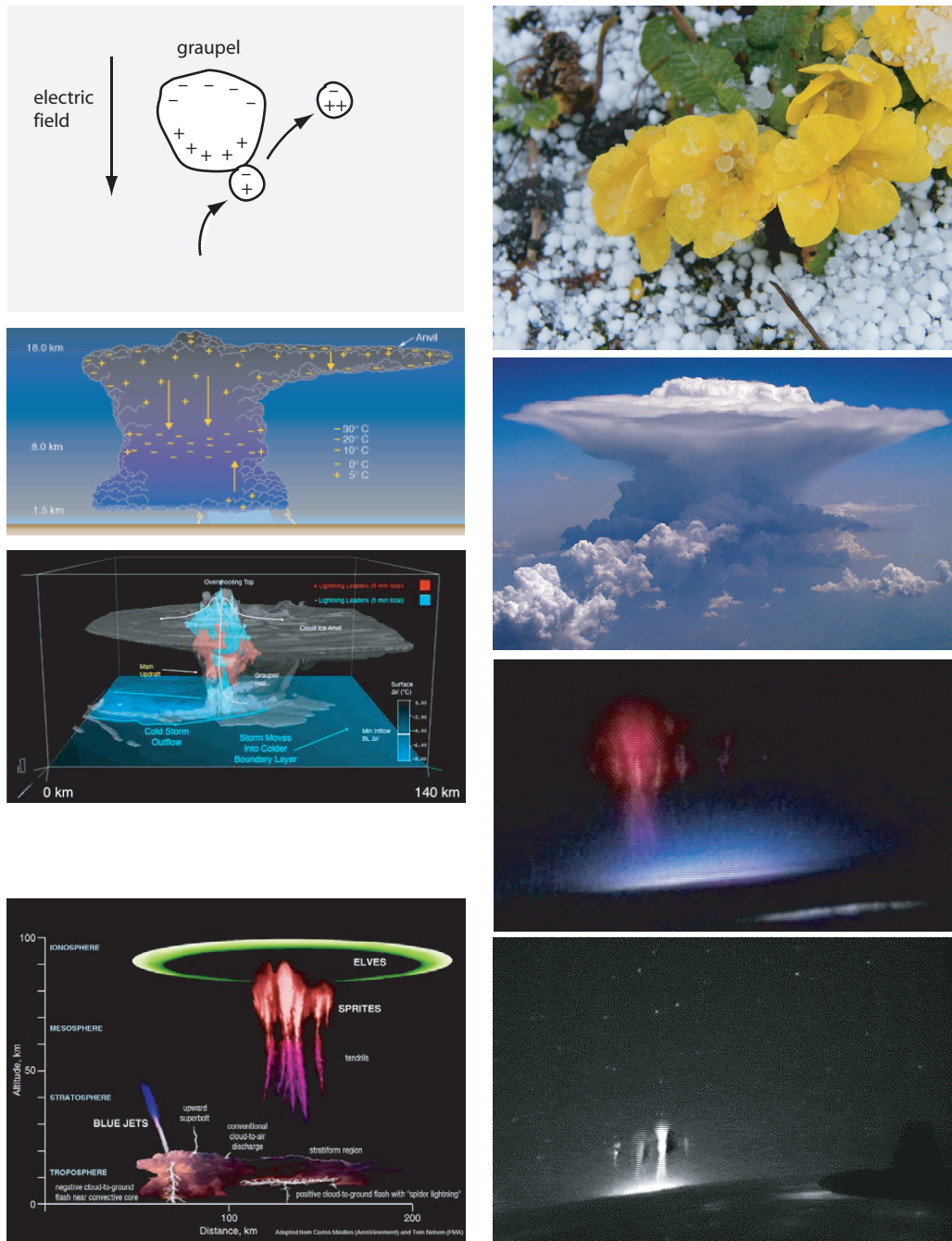
FIGURE 105 The charging and discharging of clouds: the most probable microscopic mechanism, namely charging of graupel particles by collision with ice particles, the cloud charge distribution, the three-dimensional structure and the large scale processes discovered in the past decades from airplanes (© nordique, NASA, NOAA).

other words, contrary to what one may think, lightning does not discharge the ground,
it actually charges it up!* Of course, lightning does discharge the cloud to ground po-
tential difference; but by doing so, it actually sends a negative charge down to the Earth
as a whole. Thunderclouds are batteries; the energy from the batteries comes from the
thermal uplifts mentioned above, which transport charge *against* the global ambient elec-
trical field.

Using a few electrical measurement stations that measure the variations of the electri-
cal field of the Earth it is possible to locate the position of all the lightning that comes
down towards the Earth at a given moment. Present research also aims at measuring the
activity of the related electrical sprites and elves in this way.

The ions in air play a role in the charging of thunderclouds via the charging of ice
crystals and rain drops. In general, all small particles in the air are electrically charged.
When aeroplanes and helicopters fly, they usually hit more particles of one charge than
of the other. As a result, aeroplanes and helicopters are charged up during flight. When
a helicopter is used to rescue people from a raft in high seas, the rope pulling the people
upwards must first be earthed by hanging it in the water; if this is not done, the people
on the raft could die from an electrical shock when they touch the rope, as has happened
a few times in the past.

The charges in the atmosphere have many other effects. Recent experiments have con-
firmed what was predicted back in the early twentieth century: lightning emits X-rays.
The confirmation is not easy though; it is necessary to put a detector near the lightning
flash. To achieve this, the lightning has to be directed into a given region. This is possible
using a missile pulling a metal wire, the other end of which is attached to the ground.
These experimental results are now being collated into a new description of lightning
which also explains the red-blue sprites above thunderclouds. In particular, the processes
also imply that inside clouds, electrons can be accelerated up to energies of a few MeV.

Why are sparks and lightning blue? This turns out to be a material property: the colour
comes from the material that happens to be excited by the energy of the discharge, usually
air. This excitation is due to the temperature of 30 kK inside the channel of a typical
lightning flash. For everyday sparks, the temperature is much lower. Depending on the
situation, the colour may arise from the gas between the two electrodes, such as oxygen
or nitrogen, or it may due to the material evaporated from the electrodes by the discharge.
For an explanation of such colours, as for the explanation of all colours due to materials,
we need to wait for the next part of our walk, on quantum theory.

But not only electric fields are dangerous. Also time-varying electromagnetic fields
can be. In 1997, in beautiful calm weather, a Dutch hot air balloon approached the power-
ful radio transmitter in Hilversum. After travelling for a few minutes near to the antenna,
the gondola suddenly detached from the balloon, killing all the passengers inside.

An investigation team reconstructed the facts a few weeks later. In modern gas bal-
loons the gondola is suspended by high quality nylon ropes. To avoid damage by light-
ning and in order to avoid electrostatic charging problems all these nylon ropes contain
thin metal wires which form a large equipotential surface around the whole balloon. Un-
fortunately, in the face of the radio transmitter, these thin metal wires absorbed radio
energy from the transmitter, became red hot, and melted the nylon wires. It was the first

Ref. 117

Ref. 118

Ref. 119

Ref. 120

Challenge 173 s   * The Earth is thus charged to about -1 MC. Can you confirm this?
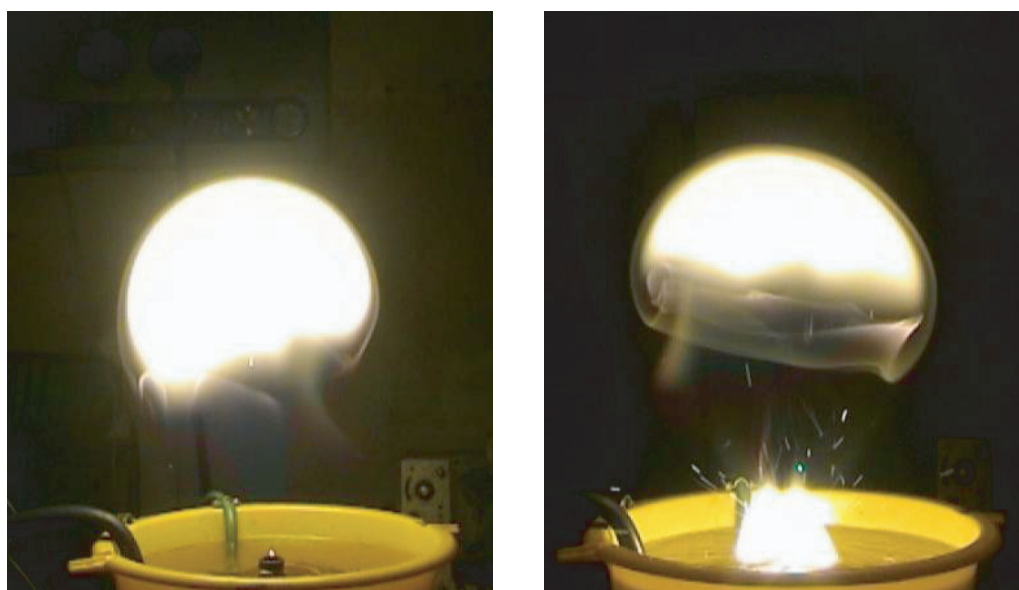
FIGURE 106 A floating plasma cloud produced in the laboratory (© Sergei Emelin and Alexei Pirozerski).

time that this had ever been observed.

Does ball lightning exist?

For hundreds of years, people have reported sightings of so-called *ball lightning*. Usually they were noticed during thunderstorms, often after a lightning had struck. With a few exceptions, nobody took these reports seriously, because no reproducible data existed.

When microwave ovens become popular, several methods to produce ball-shaped discharges became known. To observe one, just stick a toothpick into a candle, light the toothpick, and put it into (somebody else's) microwave oven at maximum power. This set-up produces a beautiful ball-like discharge. However, humans do not live in a microwave oven; therefore, this mechanism is not related to ball lightning.

The experimental situation changed completely in the years 1999 to 2001. In those years the Russian physicists Anton Egorov and Gennady Shabanov discovered a way to produce plasma clouds, or *plasmoids*, floating in air, using three main ingredients: water, metal and high voltage. If high voltage is applied to submerged metal electrodes of the proper shape and make, plasma clouds emerge from the water, about 10 to 20 cm in size, float above the surface, and disappear after about half a second. Two examples can be seen in Figure 106.

The phenomenon of floating plasmoids is still being studied. There are variations in shape, colour, size and lifetime. The spectrum of observations and techniques will surely evolve in the coming years.

An even more astonishing effect was published in 2007. A Brazilian research team found a way to make golf-ball sized discharges that seem to roll along the floor for as long as 8 s. Their method was beautifully simple: with the help of a 25 V power supply, they passed a current of 140 A through an arc at the surface of a silicon wafer. They

Ref. 121

Ref. 122

Ref. 123

discovered that small silicon particles detach and move away, while being surrounded by a luminous glow. These luminous clouds can wander around the table and floor of the laboratory, until they extinguish.

It seems that these phenomena could explain a number of ball lightning observations. But it is equally possible that other effects will still be discovered.

DOES GRAVITY MAKE CHARGES RADIATE?

We learned in the section on general relativity that gravitation has the same effects as acceleration. This means that a charge kept fixed at a certain height is equivalent to a charge accelerated by $9.8 \, \text{m/s}^2$, which would imply that it radiates electromagnetically, since all accelerated charges radiate. However, the world around us is full of charges at fixed heights, and there is no such radiation. How is this possible?

Ref. 124   The question has been a pet topic for many years. Generally speaking, the concept of radiation is not observer invariant: If one observer detects radiation, a second one does not necessarily do so as well. The exact way a radiation field changes from one observer to the other depends on the type of relative motion and on the field itself.

A detailed exploration of the problem shows that for a uniformly accelerated charge, an observer undergoing the same acceleration only detects an electrostatic field. In con-
Ref. 125   trast, an inertial observer detects a radiation field. Since gravity is (to a high precision) equivalent to uniform acceleration, we get a simple result: gravity does not make electrical charges radiate for an observer at rest with respect to the charge, as is observed. The results holds true also in the quantum theoretical description.

PLANETARY MAGNETIC FIELDS

The classical description of electrodynamics is coherent and complete; nevertheless there are still many subjects of research. Here are a few of them.

The origin of the magnetic field of the Earth, the other planets, the Sun and even of the galaxy is a fascinating topic. The way that the convection of fluids inside the planets generates magnetic fields, an intrinsically three-dimensional problem, the influence of turbulence, of non-linearities and of chaos makes it a surprisingly complex question.

The details of the generation of the magnetic field of the Earth, usually called the *geodynamo*, began to appear only in the second half of the twentieth century, when the knowl-
Ref. 25   edge of the Earth's interior reached a sufficient level. The Earth's interior starts below the Earth's crust. The *crust* is typically 30 to 40 km thick (under the continents), though it is thicker under high mountains and thinner near volcanoes or under the oceans. As already mentioned, the crust consists of large segments, the *plates*, that move with respect to one other. The Earth's interior is divided into the *mantle* – the first 2900 km from the surface – and the *core*. The core is made up of a liquid *outer* core, 2210 km thick, and a solid *inner* core of 1280 km radius. (The temperature of the core is not well known; it
Challenge 174 d   is believed to be 6 to 7 kK. Can you find a way to determine it? The temperature might have decreased a few hundred kelvin during the last 3000 million years.)

The Earth's core consists mainly of iron that has been collected from the asteroids that collided with the Earth during its youth. It seems that the liquid and electrically conducting outer core acts as a dynamo that keeps the magnetic field going. The magnetic energy comes from the kinetic energy of the outer core, which rotates with respect to the Earth's

ocean
crust
mantle
liquid core
solid core
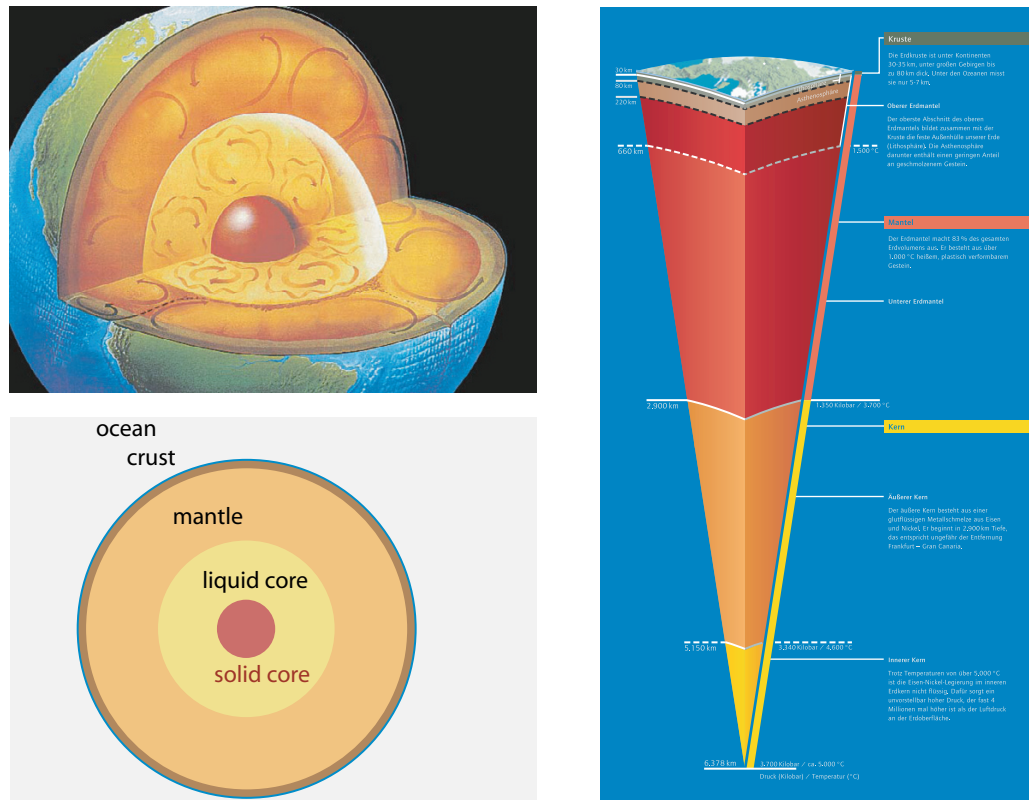
FIGURE 107 The structure of our planet (© MPI-Chemie, Mainz/GEO).

surface; the fluid can act as a dynamo because, apart from rotating, it also *convects* from deep inside the Earth to more shallow depths, driven by the temperature gradients between the hot inner core and the cooler mantle. Huge electric currents flow in complex ways through these liquid layers, maintained by friction, and create the magnetic field. Why this field switches orientation at irregular intervals of between a few tens of thousands and a few million years, is one of the central questions. The answers are difficult; experiments are not yet possible, 150 years of measurements is a short time when compared with the last transition – about 730 000 years ago – and computer simulations are extremely involved. Since the field measurements started, the dipole moment of the magnetic field has steadily diminished, presently by 5% a year, and the quadrupole moment has steadily increased. Maybe we are heading towards a surprise.* (By the way, the study of *galactic* magnetic fields is even more complex, and still in its infancy.)

LEVITATION

We have seen that it is possible to move certain objects without touching them, using a magnetic or electric field or, of course, using gravity. Is it also possible, without touching an object, to keep it fixed, floating in mid-air? Does this type of rest exist?

Ref. 126

* In 2005, it has been reported that the inner core of the Earth seems to rotate faster than the Earth's crust by up to half a degree per year.

It turns out that there are several methods of levitating objects. These are commonly divided into two groups: those that consume energy and those who do not. Among the methods that consume energy is the floating of objects on a jet of air or of water, the floating of objects through sound waves, e.g. on top of a siren, or through a laser beam coming from below, and the floating of conducting material, even of liquids, in strong radio-frequency fields. Levitation of liquids or solids by strong ultrasound waves is presently

becoming popular in laboratories. All these methods give *stationary* levitation. Another group of energy consuming methods sense the way a body is falling and kick it up again in the right way via a feedback loop; these methods are *non-stationary* and usually use magnetic fields to keep the objects from falling. The magnetic train being built in Shang-

hai by a German consortium is levitated this way. The whole train, including the passengers, is levitated and then moved forward using electromagnets. It is thus possible, using magnets, to levitate many tens of tonnes of material.

For levitation methods that do *not* consume energy – all such methods are necessarily stationary – a well-known limitation can be found by studying Coulomb's 'law' of electrostatics: no static arrangement of electric fields can levitate a *charged* object in free space or in air. The same result is valid for gravitational fields and *massive* objects;[*] in other words, we cannot produce a local minimum of potential energy in the middle of a

box using electric or gravitational fields. This impossibility is called *Earnshaw's theorem*. Speaking mathematically, the solutions of the Laplace equation $\Delta \varphi = 0$, the so-called *harmonic functions*, have minima or maxima only at the border, and never inside the domain of definition. (You proved this yourself on page 150.) The theorem can also be proved by noting that given a potential minimum in free space, Gauss' theorem for a sphere around that minimum requires that a source of the field be present inside, which is in contradiction with the original assumption.

We can deduce that it is also impossible to use electric fields to levitate an electrically *neutral* body in air: the potential energy U of such a body, with volume V and dielectric constant ε, in an environment of dielectric constant ε_0, is given by

$$\frac{U}{V} = -\frac{1}{2} \left(\varepsilon - \varepsilon_0 \right) E^2 \ . \tag{79}$$

Since the electric field E never has a maximum in the absence of space charge, and since for all materials $\varepsilon > \varepsilon_0$, there cannot be a minimum of potential energy in free space for a neutral body.[**]

To sum up, using static electric or static gravitational fields it is impossible to keep an object from falling; neither quantum mechanics, which incorporates phenomena such as antimatter, nor general relativity, including phenomena such as black holes, change this basic result.

---

[*] To the disappointment of many science-fiction addicts, this would even be true if a negative mass existed. And even though gravity is not really due to a field, but to space-time curvature, the result still holds in general relativity.

[**] It is possible, however, to 'levitate' gas bubbles in liquids – 'trap' them to prevent them from rising would be a better expression – because in such a case the dielectric constant of the environment is higher than that

of the gas. Can you find a liquid–gas combination where bubbles fall instead of rise?

For static *magnetic* fields, the argument is analogous to electrical fields: the potential energy U of a magnetizable body of volume V and permeability μ in a medium with permeability μ_0 containing no current is given by

$$\frac{U}{V} = -\frac{1}{2}\left(\frac{1}{\mu} - \frac{1}{\mu_0}\right)B^2 \tag{80}$$

and due to the inequality $\Delta B^2 \geqslant 0$, isolated maxima of a static magnetic field are not possible, only isolated minima. Therefore, it is impossible to levitate paramagnetic ($\mu > \mu_o$) or ferromagnetic ($\mu \gg \mu_0$) materials such as steel, including bar magnets, which are all attracted, and not repelled to magnetic field maxima.

There are thus two ways to realize magnetic levitation: levitating a diamagnet or using a time-dependent field. Diamagnetic materials ($\mu < \mu_0$, or $\mu_r = \mu/\mu_0 < 1$) were discovered shortly after Earnshaw published his theorem, and allow circumventing it. Indeed, diamagnetic materials, such as graphite or water, can be levitated by static magnetic fields because they are attracted to magnetic field minima. In fact, it is possible to levitate magnets if one uses a combination containing diamagnets. A few cases that can easily be replicated on a kitchen table are shown in Figure 108.

Another well-known example of diamagnetic levitation is the levitation of superconductors. Indeed, superconductors, at least those of type I, are perfects diamagnets ($\mu = 0$). In some cases, superconductors can even be *suspended* in mid-air, below a magnet. Also single atoms with a magnetic moment are diamagnets; they are routinely levitated this way and have also been photographed in this state. Single neutrons, which have a magnetic dipole moment, have been kept in magnetic bottles through magnetic levitation, until they decay.

Diamagnets levitate if $\nabla B^2 > 2\mu_0 \rho g/\chi$, where ρ is the mass density of the object and $\chi = 1 - \mu/\mu_0$ its magnetic susceptibility. Since χ is typically about 10^{-5} and ρ of order $1000\,\text{kg/m}^3$, field gradients of about $1000\,\text{T}^2/\text{m}$ are needed. In other words, levitation requires fields changes of $10\,\text{T}$ over $10\,\text{cm}$, which is nowadays common for high field laboratory magnets.

Recently, scientists have levitated pieces of wood and of plastic, strawberries, water droplets, liquid helium droplets as large as 2 cm, grasshoppers, fish and frogs (all alive and without any harm) using magnetic levitation. Indeed, animals, like humans, are all made of diamagnetic material. Humans themselves have not yet been levitated, but the feat, expected to require $40\,\text{T}$ and large amounts of electrical power, is being planned and worked on. In fact, a similar feat has already been achieved: diamagnetic levitation is being explored for the levitation of passenger trains, especially in Japan, though with little commercial success.

Time-dependent electrical or magnetic fields, e.g. periodic fields, can lead to levitation in many different ways without any consumption of energy. This is one of the methods used in the magnetic bearings of turbomolecular vacuum pumps. Also single charged particles, such as ions and electrons, are now regularly levitated with Paul traps and Penning traps. The mechanical analogy is shown in Figure 109.

Figure 110 shows a toy that allows you to personally levitate a spinning top or a spinning magnetic sphere in mid-air above a ring magnet, a quite impressive demonstration
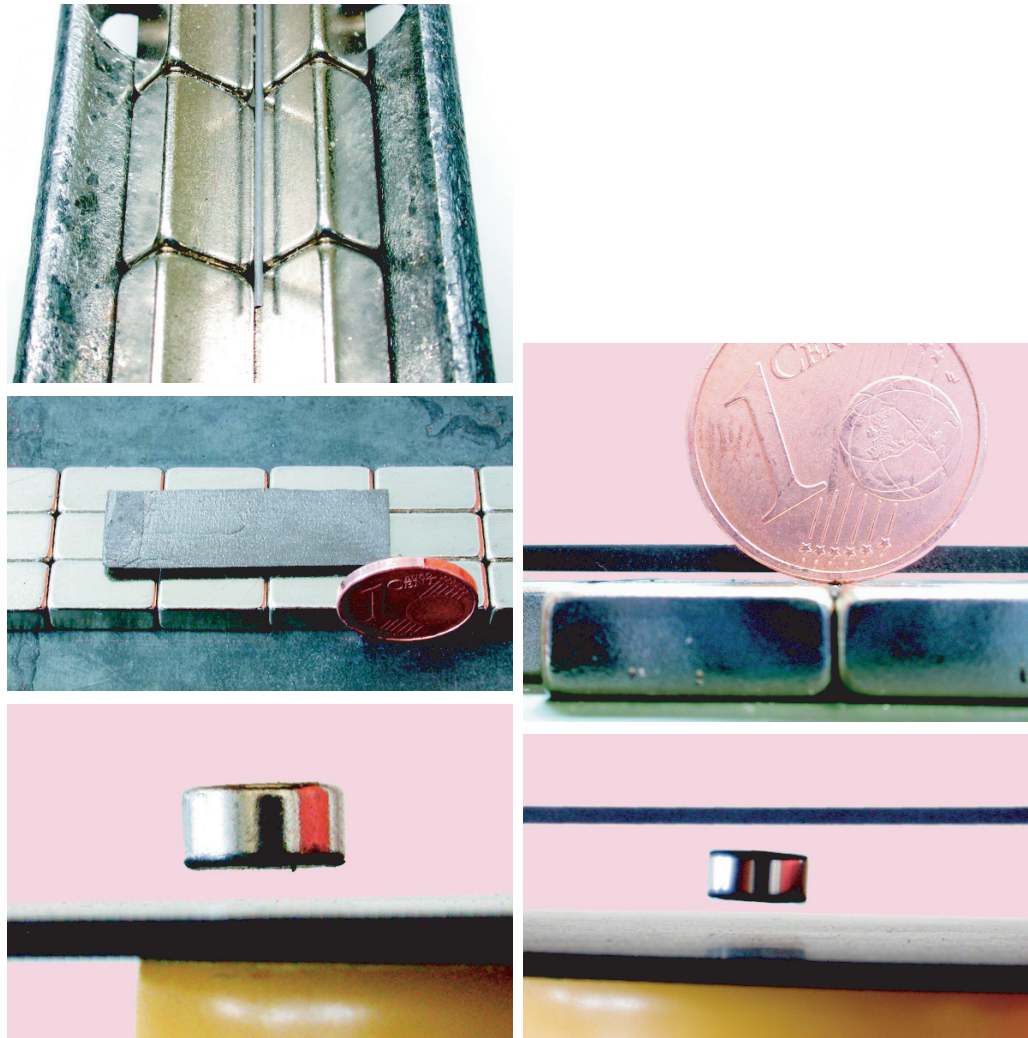
FIGURE 108 Stable diamagnetic levitation of a graphite bar over rectangular permanent magnets (above) and of two graphite plates, one seen from above and another from the side (centre); below, levitation of a 4 mm diameter NdFeB permanent magnet, above a graphite plate and between two graphite plates, near a large ring magnet (not shown) (© Joachim Schlichting from Ref. 132).

of levitation for anybody looking at it. The photo shows that is not hard to build such a device yourself.

Ref. 137

Even free electrons can be levitated, letting them float above the surface of fluid helium. In the most recent twist of the science of levitation, in 1995 Stephen Haley predicted that the suspension height of small magnetic particles above a superconducting ring should be quantized. However, the prediction has not been verified by experiment yet.

Ref. 138

For the sake of completeness we mention that nuclear forces cannot be used for levitation in everyday life, as their range is limited to a few femtometres. However, we will see later that the surface matter of the Sun is prevented from falling into the centre by

Vol. V, page 149
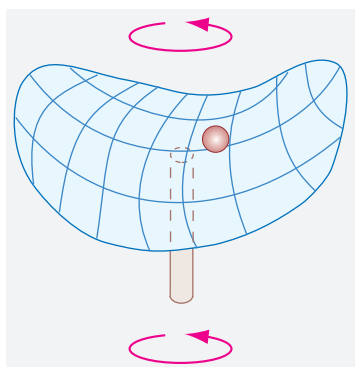
FIGURE 109 Trapping a metal sphere using a variable speed drill and a plastic saddle.
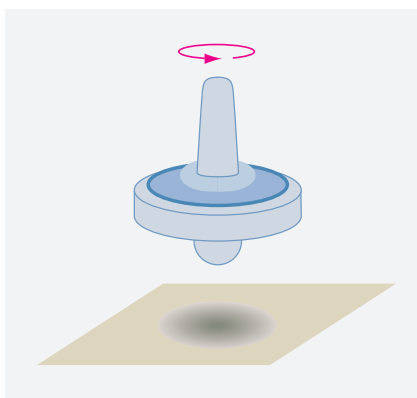
FIGURE 110 Floating 'magic' nowadays available in toy shops, left, with a spinning top and, right, with a spinning magnetic sphere levitating above a large ring magnet (© Kay Kublenz).

these interactions; we could thus say that it is indeed levitated by nuclear interactions.

Matter, levitation and electromagnetic effects

The levitation used by magicians mostly falls into another class. When David Copperfield, a magician performing for the MTV generation at the end of the twentieth century, 'flies' during his performances, he does so by being suspended on thin fishing lines that are rendered invisible by clever lighting arrangements. (How could one check this?) In fact, if we want to be precise, we should count fishing lines, plastic bags, as well as every table and chair as levitation devices. (Tabloid journalists would even call them 'anti-gravity' devices.) Contrary to our impression, a hanging or lying object is not really in contact with the suspension, if we look at the critical points with a microscope.* More about this in the quantum part of our walk.

But if this is the case, why don't we fall through a table or through the floor? We started the study of mechanics by stating that a key property of matter its *solidity*, i.e., the impossibility of having more than one body at the same place at the same time. But

Challenge 180 s

Challenge 181 ny

* The issue is far from simple: which one of the levitation methods described above is used by tables or chairs?

what is the origin of solidity? Again, we will be able to answer the question only in the forthcoming, quantum part of our adventure, but we can already collect the first clues at this point.

Solidity is due to electricity. Many experiments show that matter is constituted of charged particles; indeed, matter can be moved and influenced by electromagnetic fields in many ways. Over the years, material scientists have produced a long list of such effects, all of which are based on the existence of charged constituents. Can you find or imagine a new one? For example, can electric charge change the colour of objects?

TABLE 16 Selected matter properties related to electromagnetism, showing among other things the role it plays in the constitution of matter; at the same time a short overview of atomic, solid state, fluid and business physics.

PROPERTY	EXAMPLE	DEFINITION
thermal radiation or heat radiation or incandescence	every object	temperature-dependent radiation emitted by any macroscopic amount of matter
Interactions with charges and currents		
electrification	separating metals from insulators	spontaneous charging
triboelectricity	glass rubbed on cat fur	charging through rubbing
barometer light	mercury slipping along glass	gas discharge due to triboelectricity Ref. 140
insulation	air	no current flow below critical voltage drop
semiconductivity	diamond, silicon or gallium arsenide	current flows only when material is impure ('doped')
conductivity	copper, metals	current flows easily
superconductivity	niobium	current flows indefinitely
ionization	fire flames	current flows easily
localization (weak, Anderson)	disordered solids	resistance of disordered solids
resistivity, Joule effect	graphite	heating due to current flow
thermoelectric effects: Peltier effect, Seebeck effect, Thomson effect	ZnSb, PbTe, PbSe, BiSeTe, Bi_2Te_3, etc.	cooling due to current flow, current flow due to temperature difference, or due to temperature gradients
acoustoelectric effect	CdS	sound generation by currents, and vice versa
magnetoresistance	iron, metal multilayers	resistance changes with applied magnetic field Ref. 141
recombination	fire alarms	charge carriers combine to make neutral atoms or molecules
annihilation	positron tomography	particle and antiparticle, e.g. electron and positron, disappear into photons
Penning effect	Ne, Ar	ionization through collision with metastable atoms

TABLE 16 (Continued) Selected matter properties related to electromagnetism.

PROPERTY	EXAMPLE	DEFINITION
Richardson effect, thermal emission	BaO_2, W, Mo, used in tv and electron microscopes	emission of electrons from hot metals
skin effect	Cu	high current density on exterior of wire
pinch effect	InSb, plasmas	high current density on interior of wire
Josephson effect	Nb-Oxide-Nb	tunnel current flows through insulator between two superconductors
Sasaki–Shibuya effect	n-Ge, n-Si	anisotropy of conductivity due to applied electric field
switchable magnetism	InAs:Mn	voltage switchable magnetization Ref. 142
Interactions with magnetic fields		
Hall effect	silicon; used for magnetic field measurements	voltage perpendicular to current flow in applied magnetic field
Zeeman effect	Cd	change of emission frequency with magnetic field
Paschen–Back effect	atomic gases	change of emission frequency in strong magnetic fields
ferromagnetism	Fe, Ni, Co, Gd	spontaneous magnetization; material strongly attracted by magnetic fields
paramagnetism	Fe, Al, Mg, Mn, Cr	induced magnetization parallel to applied field; attracted by magnetic fields
diamagnetism	water, Au, graphite, NaCl	induced magnetization opposed to applied field; repelled by magnetic fields
magnetostriction	CeB_6, $CePd_2Al_3$	change of shape or volume by applied magnetic field
magnetoelastic effect	Fe, Ni	change of magnetization by tension or pressure
acoustomagnetic effect	metal alloys, anti-theft stickers	excitation of mechanical oscillations through magnetic field
spin valve effect	metal multilayers	electrical resistance depends on spin direction of electrons with respect to applied magnetic field
magneto-optical activity or Faraday effect or Faraday rotation	flint glass	polarization angle is rotated with magnetic field; different refraction index for right and left circularly polarized light, as in magneto-optic (MO) recording
magnetic circular dichroism	gases	different absorption for right- and left-circularly polarized light; essentially the same as the previous one
Majorana effect	colloids	specific magneto-optic effect

TABLE 16 (Continued) Selected matter properties related to electromagnetism.

PROPERTY	EXAMPLE	DEFINITION
photoelectromagnetic effect	InSb	current flow due to light irradiation of semiconductor in a magnetic field
Voigt effect	vapours	birefringence induced by applied magnetic field
Cotton–Mouton effect	liquids	birefringence induced by applied magnetic field
Hanle effect	Hg	change of polarization of fluorescence with magnetic field
Shubnikov–de Haas effect	Bi	periodic change of resistance with applied magnetic field
thermomagnetic effects: Ettinghausen effect, Righi–Leduc effect, Nernst effect, magneto–Seebeck effect	BiSb alloys	relation between temperature, applied fields and electric current
Ettinghausen–Nernst effect	Bi	appearance of electric field in materials with temperature gradients in magnetic fields
photonic Hall effect	CeF_3	transverse light intensity depends on the applied magnetic field Ref. 143
magnetocaloric effect	gadolinium, GdSiGe alloys	material cools when magnetic field is switched off Ref. 144
cyclotron resonance	semiconductors, metals	selective absorption of radio waves in magnetic fields
magnetoacoustic effect	semiconductors, metals	selective absorption of sound waves in magnetic fields
magnetic resonance	most materials, used for imaging in medicine for structure determination of molecules	selective absorption of radio waves in magnetic fields
magnetorheologic effect	liquids, used in advanced car suspensions	change of viscosity with applied magnetic fields
Meissner effect	type 1 superconductors, used for levitation	expulsion of magnetic field from superconductors

Interactions with electric fields

polarizability	all matter	polarization changes with applied electric field
ionization, field emission, Schottky effect	all matter, tv	charges are extracted at high fields
paraelectricity	$BaTiO_3$	applied field leads to polarization in same direction

TABLE 16 (Continued) Selected matter properties related to electromagnetism.

PROPERTY	EXAMPLE	DEFINITION
dielectricity	water	in opposite direction
ferroelectricity	$BaTiO_3$	spontaneous polarization below critical temperature
piezoelectricity	the quartz lighter used in the kitchen, human bones	polarization appears with tension, stress, or pressure
electrostriction	platinum sponges in acids	shape change with applied voltage Ref. 145
pyroelectricity	$CsNO_3$, tourmaline, crystals with polar axes; used for infrared detection	change of temperature produces charge separation
electro-osmosis or electrokinetic effect	many ionic liquids	liquid moves under applied electric field Ref. 146
electrowetting	salt solutions on gold	wetting of surface depends on applied voltage
electrolytic activity	sulphuric acid	charge transport through liquid
liquid crystal effect	watch displays	molecules turn with applied electric field
electro-optical activity: Kerr effect, Pockels effect	liquids (e.g. oil), crystalline solids	material in electric field rotates light polarization, i.e., produces birefringence
Freederichsz effect, Schadt–Helfrichs effect	nematic liquid crystals	electrically induced birefringence
Stark effect	hydrogen, mercury	colour change of emitted light in electric field
field ionization	helium near tungsten tips in field ion microscope	ionization of gas atoms in strong electric fields
Zener effect	Si	energy-free transfer of electrons into conduction band at high fields
field evaporation	W	evaporation under strong applied electric fields

Interactions with light

absorption	coal, graphite	transformation of light into heat or other energy forms (which ones?)Challenge 183 s
blackness	coal, graphite	complete absorption in visible range
colour, metallic shine	ruby	absorption depending on light frequency
photostriction	PbLaZrTi	light induced piezoelectricity
photography	AgBr, AgI	light precipitates metallic silver
photoelectricity, photoeffect	Cs	current flows into vacuum due to light irradiation
internal photoelectric effect	Si p–n junctions, solar cells	voltage generation and current flow due to light irradiation

TABLE 16 (Continued) Selected matter properties related to electromagnetism.

PROPERTY	EXAMPLE	DEFINITION
photon drag effect	p-Ge	current induced by photon momentum
emissivity	all bodies	ability to emit light
transparency	glass, quartz, diamond	low reflection, low absorption, low scattering
reflectivity	metals	light bounces on surface
polarization	pulled polymer sheets	light transmission depending on polarization angle
optical activity	sugar dissolved in water, quartz	rotation of polarization
birefringence	calcite, cornea	refraction index depends on polarization direction, light beams are split into two beams
dichroism	aminoacids, andalusite	absorption depends on circular polarization
optically induced anisotropy, Weigert effect	AgCl	optically induced birefringence and dichroism
second harmonic generation	$LiNbO_3$, KPO_4	light partially transformed to double frequency
luminescence: general term for opposite of incandescence	GaAs, tv	cold light emission
fluorescence	CaF_2, X-ray production, light tubes, cathode ray tubes	light emission during and after light absorption or other energy input
phosphorescence	$TbCl_3$	light emission due to light, electrical or chemical energy input, continuing *long after* stimulation
electroluminescence	ZnS	emission of light due to alternating electrical field
photoluminescence	ZnS : Cu, $SrAlO_4$: Eu, Dy, hyamine	light emission triggered by UV light, used in safety signs
chemoluminescence	H_2O_2, phenyl oxalate ester, dye	cold light emission used in light sticks for divers and fun
bioluminescence	glow-worm, deep sea fish	cold light emission in animals
triboluminescence	sugar	light emission during friction or crushing
thermoluminescence	quartz, feldspar	light emission during heating, used e.g. for archaeological dating of pottery Ref. 147
bremsstrahlung	X-ray generation	radiation emission through fast deceleration of electrons

TABLE 16 (Continued) Selected matter properties related to electromagnetism.

PROPERTY	EXAMPLE	DEFINITION
Compton effect	momentum measurements	change of wavelength of light, esp. X-rays and gamma radiation, colliding with matter
Čerenkov effect	water, polymer particle detectors	light emission in a medium due to particles, e.g. emitted by radioactive processes, moving faster than the speed of light in that medium
transition radiation	any material	light emission due to fast particles moving from one medium to a second with different refractive index
electrochromicity	wolframates	colour change with applied electric field
scattering	gases, liquids	light changes direction
Mie scattering	dust in gases	light changes direction
Raleigh scattering	sky	light changes direction, sky is blue
Raman effect or Smekal–Raman effect	molecular gases	scattered light changes frequency
laser activity, superradiation	beer, ruby, He–Ne	emission of stimulated radiation
sonoluminescence	air in water	light emission during cavitation
gravitoluminescence	does not exist; Challenge 184 s why?	
switchable mirror	LaH	voltage controlled change from reflection to transparency Ref. 148
radiometer effect	bi-coloured windmills	irradiation turns mill (see page 102)
luminous pressure	*idem*	irradiation turns mill directly
solar sail effect	future satellites	motion due to solar wind
acoustooptic effect	$LiNbO_3$	diffraction of light by sound in transparent materials
photorefractive materials	$LiNbO_3$, GaAs, InP	light irradiation changes refractive index
Auger effect	Auger electron spectroscopy	electron emission due to atomic reorganization after ionization by X-rays
Bragg reflection	crystal structure determination	X-ray diffraction by atomic planes
Mößbauer effect	Fe, used for spectroscopy	recoil-free resonant absorption of gamma radiation
pair creation	Pb	transformation of a photon in a charged particle–antiparticle pair
photoconductivity	Se, CdS	change of resistivity with light irradiation
optoacoustic affect, photoacoustic effect	gases, solids	creation of sound due to absorption of pulsed light; used for imaging of animal and human tissue

TABLE 16 (Continued) Selected matter properties related to electromagnetism.

PROPERTY	EXAMPLE	DEFINITION
optogalvanic effect	plasmas	change of discharge current due to light irradiation
optical nonlinear effects: parametric amplification, frequency mixing, saturable absorption, n-th harmonic generation, optical Kerr effect, etc.		
phase conjugated mirror activity	gases	reflection of light with opposite phase
Material properties		
solidity, impenetrability	floors, columns, ropes, buckets	at most one object per place at a given time
Interactions with vacuum		
Casimir effect	metals	attraction of uncharged, conducting bodies

All matter properties in the list can be influenced by electric or magnetic fields or directly depend on them. This shows that the nature of all these material properties is electromagnetic. In other words, charges and their interactions are an essential and fundamental part of the structure of objects. The table shows so many different electromagnetic properties that the motion of charges inside each material must be complex indeed. Most effects are the topic of solid state physics,* fluid and plasma physics.

Solid state physics is by far the most important part of physics, when measured by the impact it has on society. Almost all effects have applications in technical products, and give employment to many people. Can you name a product or business application for any randomly chosen effect from the table?

In our mountain ascent however, we look at only one example from the above list: thermal radiation, the emission of light by hot bodies.

Earnshaw's theorem about the impossibility of a stable equilibrium for charged particles at rest implies that the charges inside matter must be *moving*. For any charged particle in motion, Maxwell's equations for the electromagnetic field show that it radiates energy by emitting electromagnetic waves. In short, classical mechanics thus predicts that matter must radiate electromagnetic energy.

Interestingly, everybody knows from experience that this is indeed the case. Hot bodies light up depending on their temperature; the working of light bulbs thus proves that metals are made of charged particles. *Incandescence*, as it is called, requires charges. Actually, *every* body emits radiation, even at room temperature. This radiation is called *thermal radiation*; at room temperature it lies in the infrared. Its intensity is rather weak in everyday life; it is given by the general expression

$$I(T) = f T^4 \; \frac{2\pi^5 k^4}{15 c^2 h^3} \quad \text{or} \quad I(T) = f \sigma T^4 \quad \text{with} \quad \sigma = 56.7 \, \text{nW/K}^4\text{m}^2 \,, \quad (81)$$

---

* Probably the best and surely the most entertaining introductory English language book on the topic is the one by NEIL ASHCROFT & DAVID MERMIN, *Solid State Physics*, Holt Rinehart & Winston, 1976.

Challenge 185 e

Ref. 149

where f is a material-, shape- and temperature-dependent factor, with a value between zero and one, and is called the *emissivity*. The constant σ is called the *Stefan–Boltzmann black body radiation constant* or *black body radiation constant*. A body whose emissivity is given by the ideal case $f = 1$ is called a *black body*, because at room temperature such a body also has an ideal absorption coefficient and thus appears black. (Can you see why?) The heat radiation such a body emits is called *black body radiation*.

By the way, which object radiates more energy: a human body or an average piece of the Sun of the same mass? Guess first!

Challenges and fun curiosities about electromagnetic effects

'Inside a conductor there is no electric field.' This statement is often found. In fact the truth is not that simple. First, a *static* field or a *static* charge on the metal surface of a body does not influence fields and charges inside it. A closed metal surface thus forms a shield against an electric field. Can you give an explanation? In fact, a tight metal layer is not required to get the effect; a cage is sufficient. One speaks of a *Faraday cage*.

The detailed mechanism allows you to answer the following question: do Faraday cages for gravity exist? Why?

For *moving* external fields or charges, the issue is more complex. Fields due to accelerated charges – radiation fields – decay exponentially through a shield. Fields due to charges moving at constant speed are strongly reduced, but do not disappear. The reduction depends on the thickness and the resistivity of the metal enclosure used. For sheet metal, the field suppression is very high; it is not necessarily high for metal sprayed plastic. Such a device will not necessarily survive a close lightning stroke.

In practice, there is no danger if lightning hits an aeroplane or a car, as long they are made of metal. (There is one film on the internet of a car hit by lightning; the driver does not even notice.) However, if your car is hit by lightning in dry weather, you should wait a few minutes before getting out of it. Can you imagine why?

Faraday cages also work the other way round. (Slowly) changing electric fields that are inside a Faraday cage are not felt outside. For this reason, radios, mobile phones and computers are surrounded by boxes made of metal or metal-sprayed plastics. The metal keeps the so-called *electromagnetic smog* to a minimum.

There are thus three reasons to surround electric appliances by a grounded shield: to protect the appliance from outside fields, to protect people and other machines from electromagnetic smog, and to protect people against the mains voltage accidentally being fed into the box (for example, when the insulation fails). In high precision experiments, these three functions can be realized by three separate cages.

For purely magnetic fields, the situation is more complex. It is quite difficult to shield the inside of a machine from outside magnetic fields. How would you do it? In practice one often uses layers of so-called *mu-metal*; can you guess what this material does?

∗∗

Researchers are trying to detect tooth decay with the help of electric currents, using the observation that healthy teeth are bad conductors, in contrast to teeth with decay. How would you make use of this effect in this case? (By the way, it might be that the totally unrelated technique of imaging with terahertz waves could yield similar results.)

Challenge 186 s
Ref. 150
Challenge 187 s

Challenge 188 s

Ref. 151

Challenge 189 ny

Ref. 152

Challenge 190 ny

* *

Human bone is piezoelectric: it produces electric signals when stressed. When we move and grow, the electric signals are used by the body to reinforce the bones in the regions that are in need. The piezoelectricity of the bones thus controls and guides their growth. This connection is also used to make fractured bones heal more rapidly: by applying *pulsed* magnetic fields to a broken bone, the healing is stimulated and accelerated. (Static magnetic fields obviously do not work for this aim.) Also teeth are piezoelectric, and the effect plays a role in their growth.

* *

In shops, one can buy piezoelectric devices – similar to a gas lighter – that are applied to mosquito bites and are said to reduce itching and even swelling. (Some product names are 'zanza click' and 'skeeter click') Can this be true?

Challenge 191 e

* *

A team of camera men in the middle of the Sahara were using battery-driven electrical equipment to make sound recordings. Whenever the microphone cable was a few tens of metres long, they also heard a 50 Hz power supply noise, even though the next power supply was thousands of kilometres away. An investigation revealed that the high voltage lines in Europe lose a considerable amount of power by irradiation; these 50 Hz waves are reflected by the ionosphere around the Earth and thus can disturb recording in the middle of the desert. Can you estimate whether this observation implies that living directly near a high voltage line is dangerous?

Challenge 192 s

* *

When solar plasma storms are seen on the Sun, astronomers first phone the electricity company. They know that about 24 to 48 hours later, the charged particles ejected by the storms will arrive on Earth, making the magnetic field on the surface fluctuate. Since power grids often have closed loops of several thousands of kilometres, additional electric currents are induced, which can make transformers in the grid overheat and then switch off. Other transformers then have to take over the additional power, which can lead to their overheating, etc. On several occasions in the past, millions of people have been left without electrical power due to solar storms. Today, the electricity companies avoid the problems by disconnecting the various grid sections, by avoiding large loops, by reducing the supply voltage to avoid saturation of the transformers and by disallowing load transfer from failed circuits to others.

* *

If the electric field is described as a sum of components of different frequencies, its so-called *Fourier components*, the amplitudes are given by

Ref. 153

$$\hat{E}(k, t) = \frac{1}{(2\pi)^3/2} \int E(x, t)e^{-ikx} \, \mathrm{d}^3 x \qquad (82)$$

and similarly for the magnetic field. It then turns out that a Lorentz invariant quantity

N, describing the energy per circular frequency ω, can be defined:

$$N = \frac{1}{8\pi} \int \frac{|\boldsymbol{E}(k, t)|^2 + |\boldsymbol{B}(k, t)|^2}{c|\boldsymbol{k}|} \, \mathrm{d}^3 k \; . \tag{83}$$

Challenge 193 s Can you guess what N is physically? (Hint: think about quantum theory.)

* *

Faraday discovered how to change magnetism into electricity, knowing that electricity could be transformed into magnetism. (The issue is subtle. Faraday's law is not the dual of Ampère's, as that would imply the use of magnetic monopoles; neither is it the reciprocal, as that would imply the displacement current. But he was looking for a link and he found a way to relate the two observations – in a novel way, as it turned out.) Faraday also discovered how to transform electricity into light and into chemistry. He then tried to
Challenge 194 ny change gravitation into electricity. But he was not successful. Why not?

* *

At high altitudes above the Earth, gases are completely ionized; no atom is neutral. One speaks of the *ionosphere*, as space is full of positive ions and free electrons. Even though both charges appear in exactly the same number, a satellite moving through the iono-
Challenge 195 s sphere acquires a negative charge. Why? How does the charging stop?

* *

A capacitor of capacity C is charged with a voltage U. The stored electrostatic energy is $E = CU^2/2$. The capacitor is then detached from the power supply and branched on to an empty capacitor of the same capacity. After a while, the voltage obviously drops to $U/2$. However, the stored energy now is $C(U/2)^2$, which is half the original value. Where did
Challenge 196 s the energy go?

* *

Challenge 197 s How can you give somebody an electrical shock using a 4.5 V battery and some wire?

* *

An old puzzle about electricity results from the equivalence of mass and energy. It is
Ref. 154 known from experiments that the size d of electrons is surely smaller than 10^{-22} m. This
Challenge 198 e means that the electric field surrounding it has an energy content E given by at least

$$E_{\text{nergy}} = \frac{1}{2}\varepsilon_0 \int E_{\text{lectric field}}^2 \, \mathrm{d}V = \frac{1}{2}\varepsilon_0 \int_d^\infty \left(\frac{1}{4\pi\varepsilon_o} \frac{q}{r^2} \right)^2 4\pi r^2 \mathrm{d}r$$

$$= \frac{q^2}{8\pi\varepsilon_o} \frac{1}{d} > 1.2 \, \mu\text{J} \; . \tag{84}$$

On the other hand, the *mass* of an electron, usually given as $511 \, \text{keV}/c^2$, corresponds to an energy of only 82 fJ, ten million times *less* than the value just calculated. In other words, classical electrodynamics has considerable difficulty describing electrons. In fact,

a consistent description of charged point particles within classical electrodynamics is impossible. This pretty topic receives only a rare – but then often passionate – interest nowadays, because the puzzle is solved in a different way in the upcoming, quantum part of our mountain ascent.

* *

Even though the golden days of materials science are over, the various electromagnetic properties of matter and their applications in devices do not seem to be completely explored yet. About once a year a new effect is discovered that merits inclusion in the list of electromagnetic matter properties of Table 16. Among others, some newer semiconductor technologies will still have an impact on electronics, such as the recent introduction of low cost light detecting integrated circuits built in CMOS (complementary metal oxide silicon) technology.

* *

The building of light sources of high quality has been a challenge for many centuries and remains one for the future. Light sources that are intense, tunable and with large coherence length or sources that emit extreme wavelengths are central to many research pursuits. As one example of many, the first X-ray lasers have recently been built; however, they are several hundred metres in size and use modified particle accelerators. The construction of compact X-ray lasers is still many years off – if it is possible at all.

* *

But maybe the biggest challenge imaginable in classical electrodynamics is to decode the currents inside the brain. Will it be possible to read our thoughts with an apparatus placed outside the head?

   One could start with a simpler challenge: Would it be possible to distinguish the thought 'yes' from the thought 'no' by measuring electrical or magnetic fields around the head? In other words, is simple mind-reading possible? The answer is yes. This has already been achieved. Even more, using brain imaging, it is already possible to distin-
guish among simple concepts that a person has in mind.
   As we have seen above, partial mind-reading is also possible already for motion-related tasks.
   The twenty-first century will surely bring many new results also for cognitive tasks. The team first performing such a feat will become instantly famous.
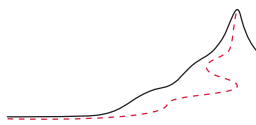
SUMMARY AND LIMITS OF CLASSICAL ELECTRODYNAMICS

Aʟʟ of classical electrodynamics can be summarized in three principles. Every adventurer should know them, because they will help us later on, when we approach the top of Motion Mountain. We will discover that we can arrive at the top only if we express things as simply as possible. The three principles of classical electrodynamics are:

> ▷ Electric charges exert forces on other charges; for charges at rest, the force falls off as the inverse square of the distance.* Equivalently, charges are surrounded by an *electromagnetic field*.

> ▷ Electric charges are conserved.

> ▷ Charges move more slowly than light.

From these three principles we can deduce all of electrodynamics. Electrodynamics is thus built on the definition of charge, the conservation of charge, and the invariance of the speed of light. In particular, we can deduce the following *basic statements*:

— The electromagnetic field is a physical observable, as shown e.g. by compass needles.
— The electromagnetic field sources are the (moving) charges, as shown by amber, lodestone or mobile phones.
— The electromagnetic field changes the motion of electrically charged objects via the Lorentz expression as, for example, shown by electric motors.
— The electromagnetic field can exist in empty space and moves in it as a wave, as shown, for example, by the stars.
— The electromagnetic field behaves like a continuous quantity and is described by Maxwell's evolution equations, as shown, for example, by mobile phones and electric toothbrushes.

More precisely, the motion of the electric field **E** and of the magnetic field **B** is described by the Lagrangian density

$$\mathcal{L} = \frac{\varepsilon_0}{2} E^2 - \frac{1}{2\mu_0} B^2 \ . \tag{85}$$

---

* Quantum theory will show that this principle, Coulomb's law, can be rephrased as: electric charges at rest emit virtual photons with a constant average rate.

Like for any motion described by a Lagrangian, the motion of the field is reversible, continuous, conserved and deterministic. However, there is quite some fun in the offing; even though this description is correct in everyday life, during the rest of our mountain ascent we will find that the last basic statement must be wrong: fields do not always follow Maxwell's equations. A simple example shows this.

At a temperature of zero kelvin, when matter does not radiate thermally, we have the paradoxical situation that the charges inside matter cannot be moving, since no emitted radiation is observed, but they cannot be at rest either, due to Earnshaw's theorem. In short, the simple existence of matter – with its charged constituents – shows that classical electrodynamics is wrong.

In fact, the overview of the numerous material properties and electromagnetic effects given in Table 16 makes the same point even more strongly; classical electrodynamics can describe many of the effects listed, *but it cannot explain the origin of any of them*. Even though few of the effects will be studied in our walk – they are not essential for our adventure – the general concepts necessary for their description will be the topic of the upcoming part of this mountain ascent, that on quantum theory.

In fact, there are *two* domains where classical electrodynamics fails.

Strong fields and gravitation

First of all, classical electrodynamics fails in regions with strong fields. When electromagnetic fields are extremely strong, their energy density will *curve* space-time. Classical electrodynamics, which assumes flat space-time, is not valid in such situations.

The failure of classical electrodynamics is most evident in the most extreme case of all: when the fields are extremely strong, they will lead to the formation of black holes. The existence of black holes, together with the discreteness of charge, imply maximum elec-
tric and magnetic field values. These upper limits were mentioned in Table 3, which lists various electric field values found in nature, and in Table 8, which lists possible magnetic field values. Can you deduce the values of these so-called *Planck fields*?

The curvature of space and electrodynamics interact in many ways. For example, the maximum force in nature limits the maximum charge that a black hole can carry. Can you find the relation? As another example, it seems that magnetic fields effectively increase the stiffness of empty space, i.e., they increase the difficulty to bend empty space. Not all interactions between gravity and electrodynamics have been studied up to now; more examples should appear in the future.

In summary, classical electrodynamics does not work for extremely high field values, when gravitation plays a role.

Charges are discrete

Classical electrodynamics fails to describe nature correctly also for a second reason, which has already been mentioned a number of times: *electric charges are discrete*. Elec- tric charges do not vary continuously, but change in fixed steps. Not only does nature show a smallest value of entropy – as we found in our exploration of heat, – and smallest amounts of matter; nature also shows a smallest charge. *Electric charge values are quantized*.

In metals, the quantization of charge is noticeable in the flow of electrons. In elec-

trolytes, i.e. electrically conducting liquids, the quantization of charge appears in the flow of charged atoms, usually called *ions*. All batteries have electrolytes inside; also water is an electrolyte, though a poorly conducting one. In plasmas, like fire or fluorescent lamps, both ions and electrons move and show the discreteness of charge. Also in radiation – from the electron beams inside TVs, channel rays formed in special glass tubes, and cosmic radiation, up to radioactivity – charges are quantized.

In all known experiments, the same smallest value e for electric charge has been found. The result is

$$e = 1.602\,176\,487(40) \times 10^{-19}\,\text{C} . \tag{86}$$

All observed electric charges in nature are multiples of this so-called *elementary charge*.

In short, like all flows in nature, also the flow of electricity is due to a flow of discrete particles. In fact, the nature of the particles differs from situation to situation: they may be electrons, ions, muons and many kind of other particles. However, the charge steps are always exactly the same. In fact, at this point of our adventure, the equality of the elementary charge for all matter particles is unexplained. We will only discover the reason only at the very end of our adventure.

Above all, a smallest charge change has a simple implication: classical electrodynamics is *wrong*. A smallest charge implies that no infinitely small test charges exist. But such infinitely small test charges are necessary to *define* electric and magnetic fields. For a *finite* test charge, the disturbance of the field introduced by the test charge itself makes a precise field measurement – and thus a precise field definition – impossible. As a consequence, the values of electric and magnetic field measured with finite test charges are always somewhat fuzzy. This fuzziness is most apparent for low field values. For example, for low intensities of light, experiments detect *photons*, discrete light particles. All low light intensities are time-averages of low photon numbers; they are not continuous fields.

The lower limit on charge magnitude also implies that there is no fully correct way of defining an instantaneous electric current in classical electrodynamics. Indeed, the position and the momentum of a charge are always somewhat fuzzy, as we will find out.

In summary, Maxwell's evolution equations are only *approximate*. Classical electromagnetism does not work for extremely low field values, when quantum effects play a role, and does not work for extremely high field values, when gravitation plays a role. These cases will be explored in the remaining legs of our adventure, those on quantum theory and those on unification. Only some effects of the discreteness of charge can be treated in classical physics. A few instructive examples follow.

How fast do charges move?

In vacuum, such as inside a colour television, charged particles accelerated by a voltage of 30 kV move with a third of the speed of light. In modern particle accelerators charges move so rapidly that their speed is indistinguishable from that of light for all practical purposes.

Inside a metal, electric signals move with speeds of the order of the speed of light. The precise value depends on the capacity and impedance of the cable and is usually in the range $0.3c$ to $0.5c$. This high speed is due to the ability of metals to easily take in arriving charges and to let others depart. The ability for rapid reaction is due to the high mobility

of the charges inside metals, which in turn is due to the small mass and size of these charges, the electrons.

The high signal speed in metals appears to contradict another determination. The *drift speed* of the electrons in a metal wire, i.e., the average speed of the charges, obviously obeys

$$v = \frac{I}{Ane} \, , \tag{87}$$

where I is the current, A the cross-section of the wire, e the charge of a single electron and n the number density of electrons. The electron density in copper is $8.5 \cdot 10^{28} \, \text{m}^{-3}$. Using a typical current of 0.5 A and a typical cross-section of a square millimetre, we get a drift speed of 0.37 μm/s. In other words, electrons move a thousand times slower than ketchup inside its bottle. Worse, if a room lamp used direct current instead of alternate current, the electrons would take several days to get from the switch to the bulb! Nevertheless, the lamp goes on or off almost immediately after the switch is activated. Similarly, the electrons from an email transported with direct current would arrive much later than a paper letter sent at the same time; nevertheless, the email arrives quickly. Are you able to explain the apparent contradiction between drift velocity and signal velocity?

Inside liquids, charges move with a different speed from that inside metals, and their charge to mass ratio is also different. We all know this from direct experience. Our *nerves* work by using electric signals and take (only) a few milliseconds to respond to a stimulus, even though they are metres long. A similar speed is observed inside semiconductors and inside batteries. In all these systems, moving charge is transported by *ions*; they are charged atoms. Ions, like atoms, are large and composed entities, in contrast to the tiny electrons.

In other matter systems, charges move both as electrons and as ions. Examples are neon lamps, fire, plasmas and the Sun. Inside atoms, electrons behave even more strangely. One tends to think that they orbit the nucleus (as we will see later) at a rather high speed, as the orbital radius is so small. However, it turns out that in most atoms many electrons do not orbit the nucleus at all. The strange story behind atoms and their structure will be told in the quantum legs of our mountain ascent.

Challenges and curiosities about charge discreteness

How would you show experimentally that electrical charge comes in smallest chunks?

* *

The discreteness of charge implies that one can estimate the size of atoms by observing galvanic deposition. How?

* *

Cosmic radiation consists of charged particles hitting the Earth. (We will discuss this in more detail later.) Astrophysicists explain that these particles are accelerated by the magnetic fields around the Galaxy. However, the expression of the Lorentz acceleration shows that magnetic fields can only change the direction of the velocity of a charge, not its magnitude. How can nature get acceleration nevertheless?

* *

What would be the potential of the Earth in volt if we could take away all the electrons of a drop of water?

* *

When a voltage is applied to a resistor, how long does it take until the end value of the current, given by Ohm's 'law', is reached? The first to answer this question was Paul Drude.[*] in the years around 1900. He reasoned that when the current is switched on, the speed v of an electron increases as $v = (eE/m)t$, where E is the electrical field, e the charge and m the mass of the electron. Drude's model assumes that the increase of electron speed stops when the electron hits an atom, loses its energy and begins to be accelerated again. Drude deduced that the average time τ up to the collision is related to the specific resistance by

$$\rho = \frac{E}{j} = \frac{E}{env} = \frac{2m}{\tau e^2 n} \; , \tag{88}$$

with n being the electron number density. The right side does not depend on E any more; it is a constant. Drude had thus explained *Ohm's relation* $U = RI$ (or $E = j\rho$) from material properties, by assuming that resistance is due to moving electrons that continuously collide and speed up again. Inserting numbers for copper ($n = 10.3 \cdot 10^{28} \,/\mathrm{m}^{-3}$ and $\rho = 0.16 \cdot 10^{-7}\,\Omega\mathrm{m}$), one gets a time $\tau = 42\,\mathrm{ps}$. This time is so short that the switch-on process can usually be neglected.

* *

Does it make sense to write Maxwell's equations *in vacuum*? Both electrical and magnetic fields require charges in order to be measured. But in vacuum there are no charges at all. And fields are defined by using infinitesimally small test charges. But, as we mentioned already, infinitesimally small charges do not exist. In fact, only quantum theory solves this issue. Are you able to imagine how?

* *

We have seen that in many cases, charge discreteness is not in contradiction with classical electrodynamics. One system that makes use of discrete charge but can nevertheless be described with classical electrodynamics merits a separate discussion: our brain.

---
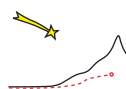
* Paul Karl Ludwig Drude (1863–1906), German physicist. A result of his electron gas model of metals was the prediction, roughly correct, that the ratio between the thermal conductivity and the electric conductivity at a given temperature should be the same for all metals. Drude also introduced c as the symbol for the speed of light.
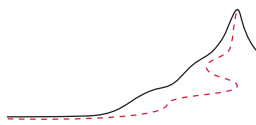
THE STORY OF THE BRAIN

> Alles was überhaupt gedacht werden kann,
> kann klar gedacht werden.*
>
> Ludwig Wittgenstein, *Tractatus*, 4.116

I N our quest for increased precision in the description of all motion around us, it s time to take a break, sit down and look back. In our walk so far, which has led us to nvestigate mechanics, general relativity and electrodynamics, we used several concepts without defining them. Examples are 'information', 'memory', 'measurement', 'set', 'number', 'infinity', 'existence', 'universe' and 'explanation'. Each of these is a common and important term. In this intermezzo, we take a look at these concepts and try to give some simple, but sufficiently precise definitions, keeping them as provocative and entertaining as possible. For example, can you explain to your parents what a concept is?

Challenge 209 e

The reason for studying definitions is simple. We need the clarifications in order to get to the top of Motion Mountain, i.e., to the full description of motion. In the past, many have lost their way because of lack of clear concepts. In order to avoid these difficulties, physics has a special guiding role. All sciences share one result: *every type of change observed in nature is a form of motion*. In this sense, but in this sense *only*, physics, focusing on motion itself, forms the basis for all the other sciences. In other words, the search for the famed 'theory of everything' is an arrogant expression for the search for a *final theory of motion*. Even though the knowledge of motion is basic, its precise description does not imply a description of 'everything': just try to solve a marriage problem using the Schrödinger equation to note the difference.

Given the basic importance of motion, it is necessary that in physics all statements on observations be as precise as possible. For this reason, many thinkers have investigated physical statements with particular care, using all criteria imaginable. Physics is detailed prattle by curious people about moving things. The criteria for precision appear once we ask: which abilities does this prattle require? You might want to fill in the list yourself before reading on.

Challenge 210 e

The abilities necessary for talking are a topic of research even today. The way that the human species acquired the ability to chat about motion is studied by evolutionary biologists. Child psychologists study how the ability develops in a single human being. Physiologists, neurologists and computer scientists are concerned with the way the brain

---

* 'Everything that can be thought at all can be thought clearly.' This and other quotes of Ludwig Wittgenstein are from the equally short and famous *Tractatus logico-philosophicus*, written in 1918, first published in 1921; it has now been translated into many other languages.
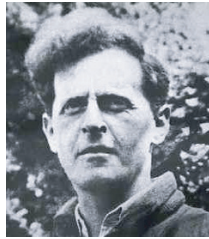
FIGURE 111 Ludwig Wittgenstein (1889–1951).

and the senses make this possible; linguists focus on the properties of the language we use, while logicians, mathematicians and philosophers of science study the general properties of statements about nature. All these fields investigate tools that are essential for the development of physics, the understanding of motion and the specification of the undefined concepts listed above. The fields structure the following exploration.

Evolution

> A hen is only an egg's way of making another egg.
>
> Samuel Butler, *Life and Habit*.

The evolution of the human species is the result of a long story that has been told in many excellent books. A summarizing table on the history of the universe was given in the exploration of general relativity. The almost incredible chain of events that has lead to one's own existence includes the formation of atoms, of the galaxies, the stars, the planets, the Moon, the atmosphere, the first cells, the water animals, the land animals, the mammals, the hominids, the humans, the ancestors, the family and finally oneself.

The way the atoms we are made of moved during this sequence, being blown through space, being collected on Earth, becoming organized to form organic matter and then people, is one of the most awe-inspiring examples of motion. Remembering and meditating about this fantastic sequence of motion every now and then can be an enriching experience.

In particular, without biological evolution, we would not be able to talk about motion; only moving bodies can study moving bodies. And without a brain, we would not be able to think or talk. Evolution was also the fount of childhood and curiosity. In this chapter we will discover that most concepts of classical physics are already introduced by every little child, in the experiences it has while growing up.

Children, laws and physics

> Physicists also have a shared reality. Other than that, there isn't really a lot of difference between being a physicist and being a schizophrenic.
>
> Richard Bandler

During childhood, everybody is a physicist. When we follow our own memories backwards in time as far as we can, we reach a certain stage, situated before birth, which forms the starting point of human experience. In that magic moment, we sensed somehow that

Ref. 159
Vol. II, page 208

Ref. 160

Challenge 211 e

Ref. 161

Ref. 162

apart from ourselves, there is something else. The first observation we make about the world, during the time in the womb, is thus the recognition that we can distinguish two parts: ourselves and the rest of the world. This distinction is an example – perhaps the first – of a large number of 'laws' of nature that we stumble upon in our lifetime. By discovering more and more distinctions we bring structure in the chaos of experience. We quickly find out that the world is made of *related parts*, such as mama, papa, milk, earth, toys, etc.

Later, when we learn to speak, we enjoy using more difficult words and we call the surroundings the *environment*. Depending on the context, we call the whole formed by oneself and the environment together the (physical) *world*, the (physical) *universe*, *nature*, or the *cosmos*. These concepts are not distinguished from each other in this walk;* they are all taken to designate the sum of all parts and their relations. They are simply taken here to designate the *whole*.

The discovery of the first distinction starts a chain of similar discoveries. We extract the numerous distinctions that are possible in the environment, in our own body and in the various types of interactions between them. The ability to distinguish is the central ability that allows us to change our view from that of the world as *chaos*, i.e., as a big mess, to that of the world as a *system*, i.e., a structured set, in which parts are related in specific ways. (If you like precision, you may ponder whether the two choices of 'chaos' and 'system' are the only possible ones.

In particular, the observation of the differences between oneself and the environment goes hand in hand with the recognition that not only are we not independent of the environment, but we are firmly tied to it in various inescapable ways: we can fall, get hurt, feel warm, cold, etc. Such relations are called *interactions*. Interactions express the observation that even though the parts of nature can be distinguished, they cannot be isolated. In other words, interactions describe the difference between the whole and the sum of its parts. No part can be defined without its relation to its environment. (Do you agree?)

Interactions are not arbitrary; just take touch, smell or sight as examples. They differ in reach, strength and consequences. We call the characteristic aspects of interactions *patterns of nature*, or *properties of nature*, or *rules of nature* or, equivalently, with their historical but unfortunate name, *'laws' of nature*. The term 'law' stresses their general validity; unfortunately, it also implies design, aim, coercion and punishment for infringement. However, no design, aim or coercion is implied in the properties of nature, nor is infringement possible. The ambiguous term 'law of nature' was made popular by René Descartes (1596–1650) and has been adopted enthusiastically because it gave weight to the laws of the state – which were far from perfect at that time – and to those of other organizations – which rarely are. The expression is an anthropomorphism coined by an authoritarian world view, suggesting that nature is 'governed'. We will therefore use the term as rarely as possible in our walk and it will, if we do, be always between 'ironical' parentheses. Nature cannot be forced in any way. The 'laws' of nature are not obligations

---

* The differences in usage can be deduced from their linguistic origins. 'World' is derived from old Germanic 'wer' – person – and 'ald' – old – and originally means 'lifetime'. 'Universe' is from the Latin, and designates the one – 'unum' – which one sees turning – 'vertere', and refers to the starred sky at night which turns around the polar star. 'Nature' comes also from the Latin, and means 'what is born'. 'Cosmos' is from Greek κόσμος and originally means 'order'.

for nature or its parts, they are obligations only for physicists and all other people: the patterns of nature oblige us to use certain descriptions and to discard others. Whenever one says that 'laws govern nature' one is talking nonsense; the correct expression is *rules describe nature*.

During childhood we learn to distinguish between interactions with the environment (or *perceptions*): some are shared with others and called *observations*, others are uniquely personal and are called *sensations*.* A still stricter criterion of 'sharedness' is used to divide the world into 'reality' and 'imagination' (or 'dreams'). Our walk will show that this distinction is not essential, provided that we stay faithful to the quest for ever increasing precision: we will find that the description of motion that we are looking for does not depend on whether the world is 'real' or 'imagined', 'personal' or 'public'.

Humans enjoy their ability to distinguish parts, which in other contexts they also call *details*, *aspects* or *entities*, and enjoy their ability to associate them or to observe the *relations* between them. Humans call this activity *classification*. Colours, shapes, objects, mother, places, people and ideas are some of the entities that humans discover first.

Our anatomy provides a handy tool to make efficient use of these discoveries: *memory*. It stores a large amount of input that is called *experience* afterwards. Memory is a tool used by both young and old children to organize their world and to achieve a certain security in the chaos of life.

Memorized classifications are called *concepts*. Jean Piaget was the first researcher to describe the influence of the environment on the concepts that a child forms. Step by step, children learn that objects are localized in space, that space has three dimensions, that objects fall, that collisions produce noise, etc. In particular, Piaget showed that space and time are not a priori concepts, but result from the interactions of every child with its environment.**

---

* A child that is unable to make this distinction among perceptions – and who is thus unable to lie – almost surely develops or already suffers from *autism*, as recent psychological research has shown.

** An overview of the origin of developmental psychology is given by J. H. FLAVELL, *The Developmental Psychology of Jean Piaget*, 1963. This work summarizes the observations by the French speaking Swiss Jean Piaget (1896–1980), the central figure in the field. He was one of the first researchers to look at child development in the same way that a physicist looks at nature: carefully observing, taking notes, making experiments, extracting hypotheses, testing them, deducing theories. His astonishingly numerous publications, based on his extensive observations, cover almost all stages of child development. His central contribution is the detailed description of the stages of development of the cognitive abilities of humans. He showed that all cognitive abilities of children, the formation of basic concepts, their way of thinking, their ability to talk, etc., result from the continuous interaction between the child and the environment.

In particular, Piaget described the way in which children first learn that they are different from the external environment, and how they then learn about the physical properties of the world. Of his many books related to physical concepts, two especially related to the topic of this walk are J. PIAGET, *Les notions de mouvement et de vitesse chez l'enfant*, Presses Universitaires de France, 1972 and *Le developpement de la notion de temps chez l'enfant*, Presses Universitaires de France, 1981, this last book being born from a suggestion by Albert Einstein. These texts should be part of the reading of every physicist and science philosopher interested in these questions.

Piaget also describes how in children the mathematical and verbal intelligence derives from sensomotorial, practical intelligence, which itself stems from habits and acquired associations to construct new concepts. Practical intelligence requires the system of reflexes provided by the anatomical and morphological structure of our organism. Thus his work shows in detail that our faculty for mathematical description of the world is based, albeit indirectly, on the physical interaction of our organism with the world.

Around the time that a child goes to school, it starts to understand the idea of *permanence of substances*, e.g. liquids, and the concept of *contrary*. Only at that stage does its subjective experience becomes *objective*, with abstract comprehension. Still later, the child's description of the world stops to be animistic: before this step, the Sun, a brook or a cloud are *alive*. In short, only after puberty does a human become ready for physics.

Even though everyone has been a physicist in their youth, most people remain *classical* physicists. In this adventure we go much further, by using all the possibilities of a toy with which nature provides us: the brain.

> “ Experience is the name everyone gives to their mistakes. ”
>
> Oscar Wilde, *Lady Windermere's Fan*.

POLYMER ELECTRONICS

The brain is electrical. This was proven in 1924, when the neurologist Hans Berger (b. 1873 Neuses, d. 1941 Jena) recorded and named the first *electroencephalogram*. In more detail, the brain is a flexible, polymer-based, metal-free, short-lived, sensitive, unreliable and electronic device. Incidentally, all these properties are shared by polymer electronics, whether alive or not. That is the reason that commercial electronics is usually silicon-based instead.

Ref. 167

The polymer electronics that forms the brain is organized like a computer. Some details are shown in Table 17, Figure 112 and Figure 113. Though the functional blocks of a brain and a computer are similar, the specific mechanisms they use are usually completely different.

WHY A BRAIN?

> “ Denken ist bereits Plastik.* ”
>
> Joseph Beuys, sculptor.

Ref. 168

Numerous observations show that sense input is processed, i.e., classified, stored and retrieved in the brain. Notably, lesions of the brain can lead to the loss of part or all of these functions. Among the important consequences of these basic abilities of the brain are thought and language. All such abilities result from the construction, from the 'hardware' of the brain.

Ref. 169

Ref. 170

Systems with the ability to deduce classifications from the input they receive are called *classifiers*, and are said to be able to *learn*. Our brain shares this property with many complex systems; the brain of many animals, but also certain computer algorithms, such as the so-called 'neural networks', are examples of such classifiers. Such systems are studied in several fields, from biology to neurology, mathematics and computer science. Classi-

Ref. 171

Ref. 172

Ref. 165   Some of his opinions on the importance of language in development are now being revised, notably through the rediscovery of the work of Lev Vigotsky, who argues that all higher mental abilities, emotions, recollective memory, rational thought, voluntary attention and self-awareness, are not innate, but learned. This learning takes place through language and culture, and in particular through the process of talking to oneself.
* 'Thinking is already sculpture.' Joseph Beuys (1920–1986).

TABLE 17 Some aspects of the human brain.

ASPECT	DETAILS	COMPUTER EQUIVALENT
Hardware		
Ultrashort term memory	5 to 9 concepts	cache
Hippocampus	novelty detector, spatial memory, learning	RAM and Flash memory
Amygdala	emotions, learning	priority scheduler of operating system
Ventral striatum, dopamine and opioid provider	rewards system	priority scheduler of operating system
Neurons in cortex	women $c.\ 19 \cdot 10^9$, men $c.\ 22 \cdot 10^9$	hard disk
Neuron number decay	women: $e^{3.05-0.00145 \cdot \text{age}/a} \cdot 10^9$, men: $e^{3.2-0.00145 \cdot \text{age}/a} \cdot 10^9$	hard disk scratching
Pulses exchanged between both brain halves	$4 \cdot 10^9/s$	internal bus speed
Synapses per neuron	10^4	
Total synapse connections	$c.\ 2 \cdot 10^{14}$	memory cells
Input pathways from the eye	$c.\ 2 \cdot 10^6$	camera wire
Input pathways from the ear	$c.\ 2 \cdot 3000$	microphone line
Input pathways from skin, mouth, and nose	$c.\ 0.5 \cdot 10^6$	sensor interfaces
Input signal capacity (total, 300 pulses/s per pathway)	$c.\ 100\,\text{MB/s}$	input bandwidth
Output pathways (muscles, organs)	$c.\ 1.5 \cdot 10^6$	actuator interfaces
Output signal capacity (total, 300 pulses/s per pathway)	$c.\ 50\,\text{MB/s}$	output bandwidth
Non-serious – probably too low – estimate of the processing capacity	10 PFlop	several dozens of supercomputers
Typical mass (Einstein's brain)	1.230 kg; varies between 0.7 and 2.0 kg	1 to 5000 kg
Power consumption (average)	1600 to 2200 kJ/d or 18 to 25 W	20 W to 20 kW
Lifetime	130 years	often only 2 years
Software and processing		
Learning	changing synapse strength through long-term potentiation	activate, classify, store
Deep sleep and learning storage	structured writing from hippocampus to cortex	clean-up and back-up to hard disk
REM (dream) sleep	offline processing	data compression in batch process
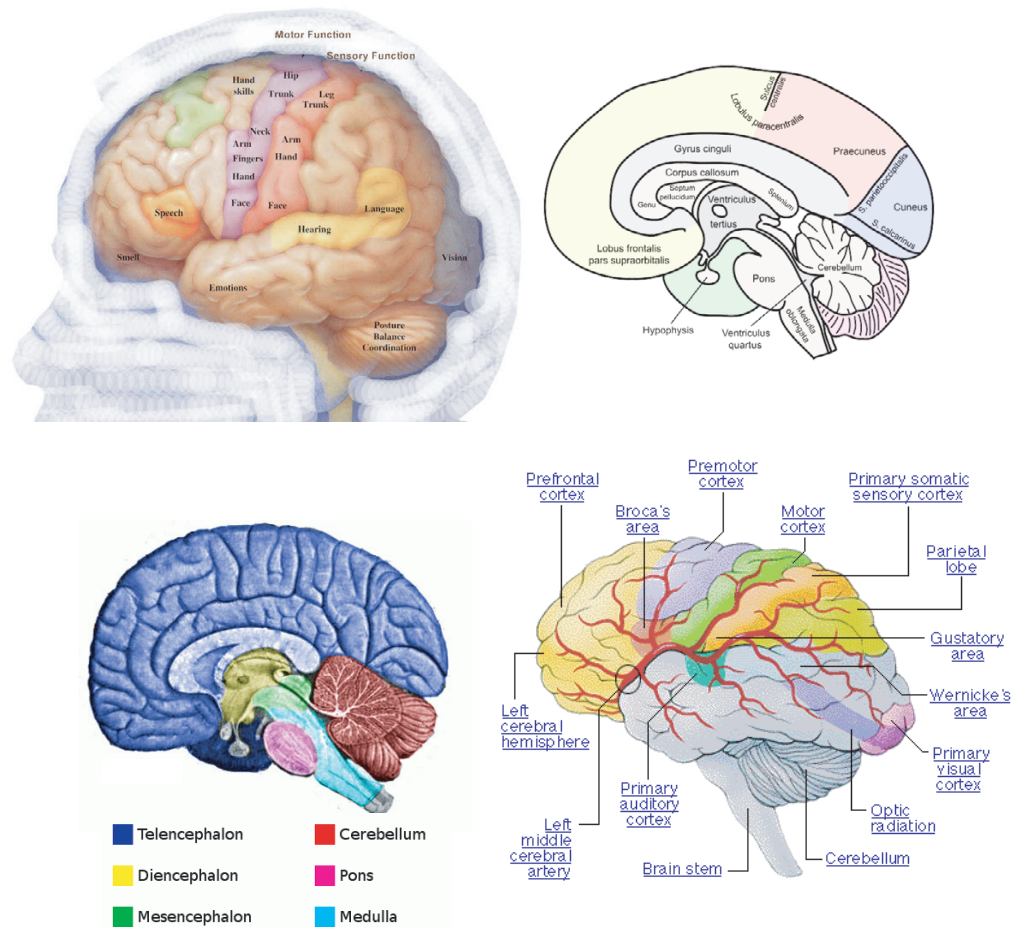
FIGURE 112 Sections and side view of the human brain, all in false colours (images WikiCommons).

fiers have the double ability to discriminate and to associate; both are fundamental to thinking.

Machine classifiers have a lot in common with the brain. As an example, following an important recent hypothesis in evolutionary biology, the necessity to cool the brain in an effective way is responsible for the upright, bipedal walk of humans. The brain needs a powerful cooling system to work well. In this it resembles modern computers, which usually have powerful fans or even water cooling systems built into them. It turns out that the human species has the most powerful cooling system of all mammals. An upright posture allowed the air to cool the body most effectively in the tropical environment where humans evolved. For even better cooling, humans have also no body hair, except on their head, where it protects the brain from direct heating by the Sun.*

All classifiers are built from smallest classifying entities, sometimes large numbers of them. Usually, the smallest units can classify input into only two different groups. The

Ref. 173

---

* The upright posture in turn allowed humans to take breath independently of their steps, a feat that many animals cannot perform. This in turn allowed humans to develop speech. Speech in turn developed the brain.
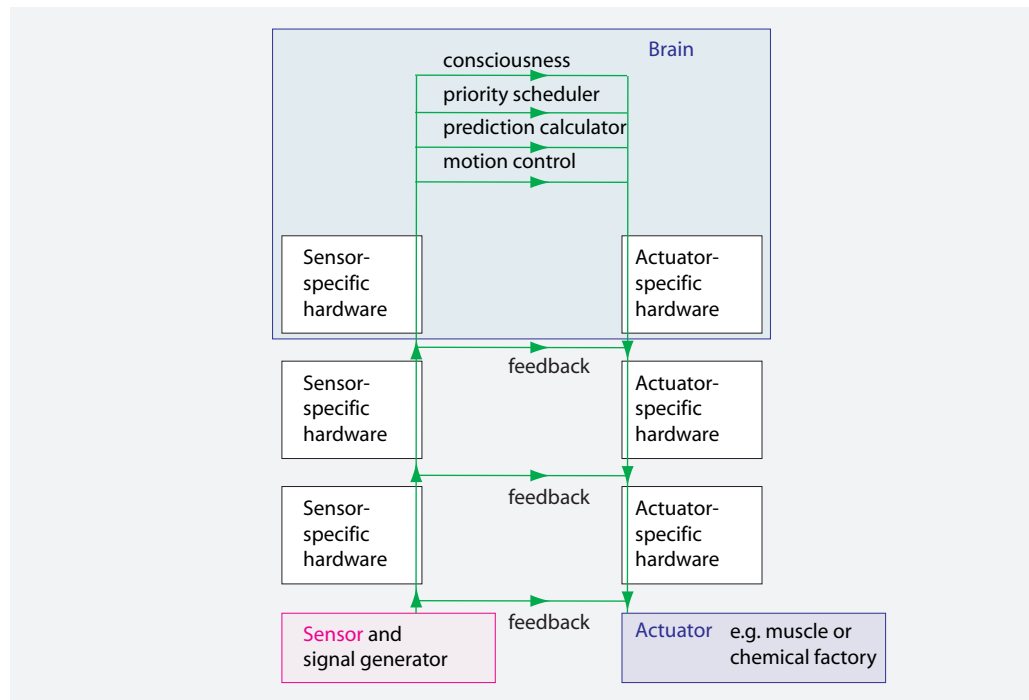
FIGURE 113 The general structure of the nervous system, with the many feedback loops it contains and the numerous sensor-specific hardware.

larger the number of these entities, often called 'neurons' by analogy to the brain, the more sophisticated classifications can be produced by the classifier. Classifiers thus work by applying more or less sophisticated combinations of 'same' and 'different'. The distinction by a child of red and blue objects is such a classification; the distinction of compact and non-compact gauge symmetry groups in quantum theory is a more elaborate classification, but relies on the same fundamental ability.

In all classifiers, the smallest classifying units interact with each other. Often these interactions are channelled via connections, and the set is then called a *network*. In these connections, signals are exchanged, via moving objects, such as electrons or photons. Thus we arrive at the conclusion that the ability of the brain to classify the physical world, for example to distinguish moving objects interacting with each other, is a consequence of the fact that it itself consists of moving objects interacting with each other. Without a powerful classifier, humans would not have become such a successful animal species. And only the motion inside our brain allows us to talk about motion in general.

Numerous researchers are identifying the parts of the brain used when different intellectual tasks are performed. The experiments become possible using magnetic resonance imaging and other methods. Other researchers are studying how thought processes can be modelled from the brain structure. Neurology is still making regular progress. In particular, it is steadily destroying the belief that thinking is *more* than a physical process. This belief results from personal fears, as you might want to test by introspection. It will disappear as time goes by. How would you argue that thought is just a physical process?

WHAT IS INFORMATION?

> These thoughts did not come in any verbal
> formulation. I rarely think in words at all. A
> thought comes, and I may try to express it in
> words afterward.
>
>   Albert Einstein

We started by stating that studying physics means to talk about motion. To talk is to transmit information. Can information be measured? Can we measure the progress of physics in this way? Is the universe made of information?

Information is the result of classification. A classification is the answer to one or to several yes–no questions. Such yes–no questions are the simplest classifications possible; they provide the basic *units* of classification, from which all others can be built. The simplest way to measure information is therefore to count the implied yes–no questions, the *bits*, leading to it. Are you able to say how many bits are necessary to define the place where you live? Obviously, the number of bits depends on the set of questions with which we start; that could be the names of all streets in a city, the set of all coordinates on the surface of the Earth, the names of all galaxies in the universe, the set of all letter combi-

nations in the address. What is the most efficient method you can think of? A variation of the combination method is used in computers. For example, the story of this walk required about a thousand million bits. But since the amount of information in a normal letter depends on the set of questions with which we start, it is impossible to define a precise measure for information in this way.

The only way to measure information precisely is to take the largest possible set of questions that can be asked about a system, and to compare it with what is known about the system. In this case, the amount of unknown information is called entropy, a concept that we have already encountered. With this approach you should able to deduce yourself

whether it is really possible to measure the advance of physics.

Since categorization is an activity of the brain and other, similar classifiers, information as defined here is a concept that applies to the result of activities by people and by other classifiers. In short, information is produced when talking about the universe – the universe itself is *not the same* as information. There is a growing number of publications based on the opposite of this view; however, this is a conceptual short circuit. Any transmission of information implies an interaction; physically speaking, this means that any information needs *energy* for transmission and *matter* for storage. Without either of these, there is no information. In other words, the universe, with its matter and energy, has to exist *before* transmission of information is possible. Saying that the universe is made of information is as meaningful as saying that it is made of toothpaste.

The aim of physics is to give a *complete* classification of all types and examples of motion, in other words, to know everything about motion. Is this possible? Or are you

able to find an argument against this endeavour?

WHAT IS MEMORY?

> The brain is my second favorite organ.
>
>   Woody Allen

Memory is the collection of records of perceptions. The production of such records is the essential aspect of observation. Records can be stored in human memory, i.e., in the brain, or in machine memory, as in computers, or in object memory, such as notes on paper. Without memory, there is no science, no life – since life is based on the records inside the DNA – and especially, no fun, as proven by the sad life of those who lose their memory.

Many animals and people have a memory, because a memory helps to move in a way that maximises reproduction and survival. Memory is found in all mammals, but also in insects and snails. The well-known sea snail *Aplysia californica* has memory – is shows conditioning, like Pawlow's dogs – even though it has only 20 000 neurons. Experiments confirm that individual memory is stored in the strength of neuron connections, the *synapses*. This statement that was made already in 1949 by the Canadian psychologist Donald Hebb. In that year Hebb specified the physical embodiment of the observations of the psychologists Sigmund Freud and William James from the 1890s, who had already deduced that memory is about the strengthening and weakening of connections inside the brain. In short, observations and learning, what we call memories, are recorded in the synapses.

Obviously every record is an object. But under which conditions does an object qualify as a record? A signature can be the record of the agreement on a commercial transaction. A single small dot of ink is not a record, because it could have appeared by mistake, for example by an accidental blot. In contrast, it is improbable that ink should fall on paper exactly in the shape of a signature. (The simple signatures of physicians are obviously exceptions.) Simply speaking, a *record* is any object, which, in order to be copied, has to be forged. More precisely, a record is an object or a situation that cannot arise nor disappear by mistake or by chance. Our personal memories, be they images or voices, have the same property; we can usually trust them, because they are so detailed that they cannot have arisen by chance or by uncontrolled processes in our brain.

Can we estimate the probability for a record to appear or disappear by chance? Yes, we can. Every record is made of a characteristic number N of small entities, for example the number of the possible ink dots on paper, the number of iron crystals in a cassette tape, the electrons in a bit of computer memory, the silver iodide grains in a photographic negative, etc. The chance disturbances in any memory are due to internal fluctuations, also called *noise*. Noise makes the record unreadable; it can be dirt on a signature, thermal magnetization changes in iron crystals, electromagnetic noise inside a solid state memory, etc. Noise is found in all classifiers, since it is inherent in all interactions and thus in all information processing.

It is a general property that internal fluctuations due to noise decrease when the size, i.e., the number of components of the record is increased. In fact, the probability p_{mis} for a misreading or miswriting of a record changes as

$$p_{mis} \sim 1/N \,, \tag{89}$$

where N is the number of particles or subsystems used for storing it. This relation appears because, for large numbers, the so-called *normal distribution* is a good approximation of almost any process. In particular, the width of the normal distribution, which determines

the probability of record errors, grows less rapidly than its integral when the number of entities is increased; for large numbers, such statements become more and more precise.

We conclude that any good record must be made from a *large* number of entities. The larger the number, the less sensitive the memory is to fluctuations. Now, a system of large size with small fluctuations is called a *(physical) bath*. Only baths make memories possible. In other words, every record contains a bath. We conclude that any *observation* of a system is the interaction of that system with a bath. This connection will be used several times in the following, in particular in quantum theory. When a record is produced by a machine, the 'observation' is usually called a *(generalized) measurement*. Are you able to specify the bath in the case of a person looking at a landscape?

From the preceding discussion we can deduce a powerful conclusion: since we have such a good memory at our disposition, we can deduce that we are made of many small parts. And since records exist, the world must also be made of a large number of small parts. No microscope of any kind is needed to confirm the existence of molecules or similar small entities; such a tool is only needed to determine the *sizes* of these particles. Their existence can be deduced simply from the observation that we have memory. (Of course, another argument proving that matter is made of small parts is the ubiquity of noise.)

A second conclusion was popularized in the late 1920s by Leo Szilard. Writing a memory does not produce entropy; it is possible to store information into a memory without increasing entropy. However, entropy is produced in every case that the memory is *erased*. It turns out that the (minimum) entropy created by erasing one bit is given by

$$S_{\text{per erased bit}} = k \ln 2 \ , \tag{90}$$

and the number $\ln 2 \approx 0.69$ is the natural logarithm of 2. Erasing thus on the one hand reduces the disorder of the data – the local entropy–, but on the other hand increases the total entropy. As is well known, energy is needed to reduce the entropy of a local system. In short, any system that erases memory requires energy. For example, a logical AND gate effectively erases one bit per operation. Logical thinking thus requires energy. It is also known that *dreaming* is connected with the erasing and reorganization of information. Could that be the reason that, when we are very tired, without any energy left, we do not dream as much as usual?

Entropy is thus necessarily created when we forget. This is evident when we remind ourselves that forgetting is similar to the deterioration of an ancient manuscript. Entropy increases when the manuscript is not readable any more, since the process is irreversible and dissipative.* Another way to see this is to recognize that to clear a memory, e.g. a magnetic tape, we have to put energy into it, and thus increase its entropy. Conversely, *writing* into a memory can often *reduce* entropy; we remember that signals, the entities

---

* As Wojciech Zurek clearly explains, the entropy created inside the memory is the main reason that even Maxwell's demon cannot reduce the entropy of two volumes of gases by opening a door between them in such a way that fast molecules accumulate on one side and slow molecules accumulate on the other. (Maxwell had introduced the 'demon' in 1871, to clarify the limits posed by nature to the gods.) This is just another way to rephrase the old result of Leo Szilard, who showed that the measurements by the demon create more entropy than they can save. And every measurement apparatus contains a memory.

To play being Maxwell's demon, click on the www.wolfenet.com/~zeppelin/maxwell.htm website.

that write memories, carry negative entropy. For example, the writing of magnetic tapes usually reduces their entropy.

THE CAPACITY OF THE BRAIN

> " Computers are boring. They can give only answers.
>
> (Wrongly) attributed to Pablo Picasso "

The human brain is built in such a way that its fluctuations cannot destroy its contents. The brain is well protected by the skull for exactly this reason. In addition, the brain literally grows connections, called *synapses*, between its various *neurons*, which are the cells doing the signal processing. The neuron is the basic processing element of the brain, performing the basic classification. It can only do two things: to fire and not to fire. (It is possible that the time at which a neuron fires also carries information; this question is not yet settled.) The neuron fires depending on its input, which comes via the synapses from hundreds of other neurons. A neuron is thus an element that can distinguish the inputs it receives into two cases: those leading to firing and those that do not. Neurons are thus classifiers of the simplest type, able only to distinguish between two situations.

Ref. 179, Ref. 180

Every time we store something in our long term memory, such as a phone number, the connection strength of existing synapses is changed or new synapses are grown. The connections between the neurons are much stronger than the fluctuations in the brain. Only strong disturbances, such as a blocked blood vessel or a brain lesion, can destroy neurons and lead to loss of memory.

As a whole, the brain provides an extremely efficient memory. Despite intense efforts, engineers have not yet been able to build a memory with the capacity of the brain in the same volume. Let us estimated this memory capacity. By multiplying the number of neurons, about 10^{11},[*] by the average number of synapses per neuron, about 100, and also by the estimated average number of bits stored in every synapse, about 10[**], we arrive at a conservative estimate for the storage capacity of the brain of about

$$M_{\text{rewritable}} \approx 10^{14} \, \text{bit} \approx 10^4 \, \text{GB} .\tag{91}$$

(One *byte*, abbreviated B, is the usual name for eight bits of information.) Note that evolution has managed to put as many neurons in the brain as there are stars in the galaxy, and that if we add all the synapse lengths, we get a total length of about 10^{11} m, which corresponds to the distance to from the Earth to the Sun. Our brain truly is *astronomically* complex.

However, this standard estimate of 10^{14} bits is not really correct! It assumes that the only component storing information in the brain is the synapse strength. Therefore it only measures the *erasable* storage capacity of the brain. In fact, information is also stored in the structure of the brain, i.e., in the exact configuration in which every cell is connected to other cells. Most of this structure is fixed at the age of about two years, but it continues to develop at a lower level for the rest of human life. Assuming that for each

* The number of neurons seems to be constant, and fixed at birth. The growth of interconnections is highest between age one and three, when it is said to reach up to 10^7 new connections per second.
** This is an average. Some types of synapses in the brain are known to store only one bit.

of the N cells with n connections there are $f\,n$ connection possibilities, this *write once* capacity of the brain can be estimated as roughly $N\sqrt{f n}\,f n \log f n$ bits. For $N = 10^{11}$, $n = 10^2$, $f = 6$, this gives

Challenge 222 e

$$M_{\text{writeonce}} \approx 10^{16}\,\text{bit} \approx 10^6\,\text{GB} . \tag{92}$$

Ref. 181   Structural brain changes are measurable. Recent measurements confirmed that bilingual persons, especially early bilinguals, have a higher density of grey mass in the small parietal cortex on the left hemisphere of the brain. This is a region mainly concerned with language processing. The brain thus makes also use of structural changes for optimized storage and processing. Similar structure changes are also known for other populations, such as autistics, homophiles and hyperactive children. Intense and prolonged experiences during pregnancy or childhood seem to induce such structural developments.

Sometimes it is claimed that people use only between 5% or 10% of their brain capacity. This myth, which goes back to the nineteenth century, would imply that it is possible to measure the actual data stored in the brain and compare it with its capacity to an impossible accuracy. Alternatively, the myth implies that the processing capacity can be measured. It also implies that nature would develop and maintain an organ with 90 % overcapacity, wasting all the energy and material to build, repair and maintain it. The myth is wrong.

The large storage capacity of the brain also shows that human memory is filled by the environment and is not inborn: one human ovule plus one sperm have a mass of about 1 mg, which corresponds to about $3 \cdot 10^{16}$ atoms. Obviously, fluctuations make it impossible to store 10^{16} bits in these systems. In fact, nature stores only about $6 \cdot 10^9$ DNA base pairs or $12 \cdot 10^9$ bits in the genes of a fecundated ovule, using $3 \cdot 10^6$ atoms per bit. In contrast, a typical brain has a mass of 1.5 to 2 kg and contains about 5 to $7 \cdot 10^{25}$ atoms, which makes it as efficient a memory as an ovule. The difference between the number of bits in human DNA and those in the brain nicely shows that almost all information stored in the brain is taken from the environment; it cannot be of genetic origin, even allowing for smart decompression of stored information.

In total, all the tricks used by nature result in the most powerful classifier yet known.[*] Are there any limits to the brain's capacity to memorize and to classify? With the tools that humans have developed to expand the possibilities of the brain, such as paper, writing and printing to help memory, and the numerous tools available to simplify and to abbreviate classifications explored by mathematicians, brain classification is only limited

Ref. 182   by the time spent practising it. Without tools, there are strict limits, of course. The two-millimetre thick cerebral cortex of humans has a surface of about four sheets of A4 paper, a chimpanzee's yields one sheet and a monkey's is the size of a postcard. It is estimated that the total intellectually accessible memory is of the order of

$$M_{\text{intellectual}} \approx 1\,\text{GB} , \tag{93}$$

though with a large experimental error.

---

[*] Also the power consumption of the brain is important: even though it contains only about 2% of the body's mass, is uses 25% of the energy taken in by food.

The brain is also unparalleled in its processing capacity. This is most clearly demonstrated by the most important consequence deriving from memory and classification: thought and language. Indeed, the many types of thinking or language we use, such as comparing, distinguishing, remembering, recognizing, connecting, describing, deducing, explaining, imagining, etc., all describe different ways to classify memories or perceptions. In the end, every type of thinking or talking directly or indirectly classifies observations. But how far are computers from achieving this! The first attempt, in 1966, was a programming joke by Joseph Weizenbaum: the famous chatterbot program Eliza (try it at www.manifestation.com/neurotoys/eliza.php3) is a parody of a psychoanalyst. Even today, over 40 years later, conversation with a computer program, such as Friendbot (found at www.friendbot.co.uk), is still a disappointing experience. We need to understand the reasons for this slow development.

Incidentally, even though the brains of sperm whales and of elephants can be five to six times as heavy as those of humans, the number of neurons and connections, and thus the capacity, is lower than for humans. Snails, ants, small fish have neuron numbers of the order of 10 000; the well-studied nematode *Caenorhabditis elegans* has only 302, though other animals have even fewer.

Curiosities about the brain

Teachers should all be brain experts. The brain learns best when it has an *aim*. Without an aim, both the lecture preparation and the lecture performance will loose most of its possible effects. How many teachers state the aim of their class at its beginning?

The brain also leans best when it is *motivated*. Different students need different motivations: potential applications, curiosity, competition, activation of already acquired knowledge, impressing the opposite sex, or exploring the unknown. And students need motivations on different levels of difficulty. Which teacher provides this mix?

Finally, brains in students have different ways to *create concepts*: using words, sounds, images, emotions, body sensations, etc. Which teacher addresses them all in his lessons?

∗ ∗

The brain plays strange games on the people that carry it. Modern research has shown that school pupils can be distinguished into five separate groups.

1. Smart students
2. Uninterested students
3. Students that overestimate themselves (often, but not always, boys)
4. Students that underestimate themselves (often, but not always, girls)
5. Struggling/weak students

This has to be kept in mind when teaching classes. To which group do/did you belong?

∗ ∗

Many cognitive activities of the brain are located in specific regions of the *cerebral cortex*, also called *grey matter* (see Figure 112). It is known that all grey matter is built of a large number of parallel, but largely independent structures, the so-called *neocortical column*s; they are similar to microprocessors. Each neocortical column has input and outputs, but works independently of the others; it is about 2 mm in height, 0.5 mm in diameter, and

Ref. 183

contains about 10 00 neurons of various types. The human cortex contains several millions of these columns, arranged in six layers. At present, researchers are able to simulate *one* neocortical column with *one* supercomputer. For more details, see the bluebrain.epfl.ch website. In short, your brain corresponds to several million supercomputers. Take good care of it.

* *

The brain has many interesting sides. The technique of *neurofeedback* is an example. A few electrodes are attached to the skin of the head, and a feedback loop is created with help of a visualization on a screen. Such a visualisation helps to put oneself into high-theta state – corresponding to deep relaxation –, or into the SMR state – corresponding to rest and concentration –, or into alpha-dominated states – corresponding to relaxation with closed eyes. Learning to switch rapidly between these states is helping athletes, surgeons, dancers, musicians, singers and children with attention deficit syndrome. After a few sessions, the effects keeps for over a year. For attention deficit syndrome, the results are as good as with medication.

Ref. 184

* *

One interesting side of the human brain is the wide range of *passions* it produces. For example, there are people whose passion drives them to dedicate all their life to singing. There are people whose life-long passion is to invent languages; John Ronald Tolkien is the most famous example. There are other people whose passion is to help murderers to find peace of mind. Some people dedicate their life to raising handicapped children unwanted by their parents. Other people dedicate their life to implementing rapid solutions for infrastructure problems – water, gas and electricity supplies – in cities under war. The examples one can find are fascinating.

* *

Many functions in the brain are not performed by the programmable part of the brain, the cortex, but by specialized hardware. The list of known specialized hardware parts of the brain is still growing, as discoveries are still being made. Researchers have discovered dedicated neurons that control the walking process in each leg, dedicated neurons – the so-called *mirror neurons* – that re-enact what people we see are feeling or acting, and dedicated neurons from the eye to the brain that control the day–night cycle. These recent discoveries complement the older ones that there is specialized hardware for every sense in the neural system, from touch to smell to proprioception. In short, many basic functions of the neural system are wired in, and many advanced functions are as well. The full list of wired-in systems is not known yet. For example, only future research will help us to understand how much of our subconscious is due to hardware, and how much is due to the software in the cortex.

Ref. 185
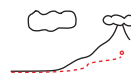
THOUGHT AND LANGUAGE

> ❝ Reserve your right to think, for even to think
> wrongly is better than not to think at all.
>
> Hypatia of Alexandria ❞

L ANGUAGE possibly is the most wonderful gift of human nature. We have all
earned it from somebody who cared about us. Nevertheless, the origins of
anguage are hidden in the distant past of humanity. But we must explore lan-
guage, because we have repeatedly stated that physics is talking about motion. Physics is
a precise language specialized for motion. We will find out in our walk that this is not a
restriction, because everything in the world moves. But our quest for precision demands
that we explore the meaning, the use and the limits of language.

What is language?

> ❝ Ein Satz kann nur sagen, *wie* ein Ding ist, nicht
> *was* es ist.*
>
> Ludwig Wittgenstein, *Tractatus*, 3.221 ❞

Using the ability to produce sounds and to put ink on paper, people attach certain *sym-
bols*,** also called *words* or *terms* in this context, to the many partitions they specify with
the help of their thinking. Such a categorization is then said to define a *concept* or *notion*,
and is set in *italic typeface* in this text. A standard set of concepts forms a language.***
Ref. 186   In other words, we have:

> ▷ A *(human) language* is a standard way of symbolic interaction between
> people.

---

* 'A proposition can only say *how* a thing is, not *what* it is.'

** A symbol is a type of *sign*, i.e., an entity associated by some convention to the object it refers. Following
Charles Peirce (1839–1914) – see www.peirce.org – the most original philosopher born in the United States,
a symbol differs from an *icon* (or *image*) and from an *index*, which are also attached to objects by convention,
in that it does not resemble the object, as does an icon, and in that it has no contact with the object, as is
the case for an index.

*** The recognition that language is based on a partition of ideas, using the various differences between
them to distinguish them from each other, goes back to the Swiss thinker Ferdinand de Saussure (1857–
1913), who is regarded as the founder of linguistics. His textbook *Cours de linguistique générale*, Editions
Payot, 1985, has been the reference work of the field for over half a century. Note that Saussure, in contrast
to Peirce, prefers the term 'sign' to 'symbol', and that his definition of the term 'sign' includes also the object
to which it refers.

TABLE 18 Language basics.

ASPECT	VALUE
Human phonemes	*c.* 70
English phonemes	44
German phonemes	40
Italian phonemes	30
Words of the English language (more than most languages, with the possible exception of German)	*c.* 350 000
Number of languages on Earth in the year 2000	*c.* 6000

There are human languages based on facial expressions, on gestures, on spoken words, on whistles, on written words, and more. The use of *spoken* language is considerably younger than the human species; it seems that it appeared only about two hundred thousand years ago. Written language is even younger, namely only about six thousand years old. But the set of concepts used, the *vocabulary*, is still expanding. For humans, the understanding of language begins soon after birth (perhaps even before), the active use begins at around a year of age, the ability to read can start as early as two, and personal vocabulary continues to grow as long as curiosity is alive.

Physics being a lazy way to chat about motion, it needs language as an essential tool. Of the many aspects of language, from literature to poetry, from jokes to military orders, from expressions of encouragement, dreams, love and emotions, physics uses only a small and rather special segment. This segment is defined by the inherent restriction to talk about motion. Since motion is an observation, i.e., an interaction with the environment that several people experience in the same way, this choice puts a number of restrictions on the contents – the vocabulary – and on the form – the grammar – of such discussions.

For example, from the definition that observations are shared by others, we get the requirement that the statements describing them must be translatable into all languages. But when can a statement be translated? On this question two extreme points of view are possible: the first maintains that *all* statements can be translated, since it follows from the properties of human languages that each of them can express every possible statement. In this view, we can say:

▷ Only sign systems that allow one to express the complete spectrum of human messages form a *human language*.

This definition of language distinguishes human spoken and sign language from animal languages, such as the signs used by apes, birds or honey bees, and also from computer languages, such as Pascal or C. With this meaning of language, all statements can be translated by definition.

It is more challenging for a discussion to follow the opposing view, namely that precise translation is possible only for those statements which use terms, word types and grammatical structures found in *all* languages. Linguistic research has invested consid-

TABLE 19 The semantic primitives, following Anna Wierzbicka.

I, you, someone, something, people	[substantives]
this, the same, one, two, all, much/many	[determiners and quantifiers]
know, want, think, feel, say	[mental predicates]
do, happen	[agent, patient]
good, bad	[evaluative]
big, small	[descriptors]
very	[intensifier]
can, if (would)	[modality, irrealis]
because	[causation]
no (not)	[negation]
when, where, after (before), under (above)	[time and place]
kind of, part of	[taxonomy, partonomy]
like	[hedge/prototype]

erable effort in the distillation of phonological, grammatical and semantic *universals*, as they are called, from the 6000 or so languages thought to exist today.*

The investigations into the *phonological* aspect, which showed for example that every language has at least two consonants and two vowels, does not provide any material for the discussion of translation.** Studying the *grammatical* (or *syntactic*) aspect, one finds that all languages use smallest elements, called 'words', which they group into sentences. They all have pronouns for the first and second person, 'I' and 'you', and always contain nouns and verbs. All languages use *subjects* and *predicates* or, as one usually says, the three entities *subject*, *verb* and *object*, though not always in this order. Just check the languages you know.

On the *semantic* aspect, the long list of lexical universals, i.e., words that appear in all languages, such as 'mother' or 'Sun', has recently been given a structure. The linguist Anna Wierzbicka performed a search for the building blocks from which all concepts can be built. She looked for the definition of every concept with the help of simpler ones, and continued doing so until a fundamental level was reached that cannot be further reduced. The set of concepts that are left over are the primitives. By repeating this exercise in many languages, Wierzbicka found that the list is the same in all cases. She thus had discovered *universal semantic primitives*. In November 1992, the list contained the terms given in Table 19.

Challenge 223 e

---

* A professional database by the linguist Merritt Ruhlen with 5700 languages and many details on each language can be found at ehl.santafe.edu/intro1.htm. A long but unprofessional list with 6 900 languages (and with 39 000 language and dialect names) can be found on the website www.ethnologue.com. Beware, it is edited by a fringe religious group that aims to increase the number of languages as much as possible.

It is estimated that 15 000 ± 5 000 languages have existed in the past.

Nevertheless, in today's world, and surely in the sciences, it is often sufficient to know one's own language plus English. Since English is the language with the largest number of words, learning it well is a greater challenge than learning most other languages.

Ref. 187

** Studies explore topics such as the observation that in many languages the word for 'little' contains an 'i' (or high pitched 'e') sound: petit, piccolo, klein, tiny, pequeño, chiisai; exceptions are: small, parvus.

Following the life-long research of Anna Wierzbicka and her research school, all these concepts exist in all languages of the world studied so far.* They have defined the meaning of each primitive in detail, performed consistency checks and eliminated alternative approaches. They have checked this list in languages from all language groups, in languages from all continents, thus showing that the result is valid everywhere. In every language all other concepts can be defined with the help of the semantic primitives.

Simply stated, learning to speak means learning these basic terms, learning how to combine them and learning the names of these composites. The definition of language given above, namely as a means of communication that allows one to express everything one wants to say, can thus be refined: a *human language* is any set of concepts that includes the universal semantic primitives.

For physicists – who aim to talk in as few words as possible – the list of semantic primitives has three facets. First, the approach is similar to physics' own aim: the idea of primitives gives a structured summary of everything that can be said, just as the atomic elements structure all objects that can be observed. Second, the list of primitives can be structured. In fact, the list of primitives can be divided into two groups: one group contains all terms describing motion (do, happen, when, where, feel, small, etc. – probably a term from the semantic field around light or colour should be added) and the other group contains all terms necessary to talk about abstract sets and relations (this, all, kind of, no, if, etc.). Even for linguistics, aspects of motion and logical concepts are the basic entities of human experience and human thinking. To bring the issue to a point, the semantic primitives contain the basic elements of physics and the basic elements of mathematics. All humans are thus both physicists and mathematicians. The third point is that the list of primitives is too long. The division of the list into two groups directly suggests shorter lists; we just have to ask physicists and mathematicians for concise summaries of their respective fields. To appreciate this aim, try to define what 'if' means, or what an 'opposite' is – and explore your own ways of reducing the list.

Reducing the list of primitives is also one of our aims in this adventure. We will explore the mathematical group of primitives in this chapter the physical group will occupy us in the rest of our adventure. However, a shorter list of primitives is not sufficient. Our goal is to arrive at a list consisting of only *one* basic concept. Reaching this goal is not simple, though. First, we need to check whether the set of classical physical concepts that we have discovered so far is complete. Can classical physical concepts describe *all* observations? The upcoming quantum part of our adventure is devoted to this question. The second task is to reduce the list. This task is not straightforward; we have already discovered that physics is based on a circular definition: in Galilean physics, space and time are defined using matter, and matter is defined using space and time. We will need quite some effort to overcome this obstacle. The final part of this text tells the precise story. After numerous adventures we will indeed discover a basic concept on which all other concepts are based.

---

* It is easy to imagine that this research steps on the toes of many people. A list that maintains that 'true', 'good', 'creation', 'life', 'mother' or 'god' are composite will elicit violent reactions, despite the correctness of the statements. Indeed, some of these terms were added in the 1996 list, which is somewhat longer. In addition, a list that maintains that we only have about thirty basic concepts in our heads is taken to be offensive by many small minds.

We can summarize all the above-mentioned results of linguistics in the following way. By constructing a statement made only of subject, verb and object, consisting only of nouns and verbs, using only concepts built from the semantic primitives, we are sure that it can be translated into all languages. This explains why physics textbooks are often so boring: the authors are often too afraid to depart from this basic scheme. On the other hand, research has shown that such straightforward statements are not restrictive: with them one can say everything that can be said.

> Jedes Wort ist ein Vorurteil.
> Friedrich Nietzsche*

> Every word was once a poem.
> Ralph Waldo Emerson**

What is a concept?

> Concepts are merely the results, rendered permanent by language, of a previous process of comparison.
> William Hamilton

There is a group of people that has taken the strict view on translation and on precision to the extreme. They build all concepts from an even smaller set of primitives, namely only two: 'set' and 'relation', and explore the various possible combinations of these two concepts, studying their classifications. Step by step, this radical group, commonly called *mathematicians*, came to define with full precision concepts such as numbers, points, curves, equations, symmetry groups and more. The construction of these concepts is summarized partly in the following and partly in the following volume of this adventure.

However, despite the involved precision, in fact precisely because of it, no mathematical concept talks about nature or about observations.*** Therefore the study of motion needs other, more useful concepts. What properties must a useful concept have? For example, what is 'passion' and what is a 'cotton bud'? Obviously, a useful concept implies a list of its parts, its aspects and their internal relations, as well as their relation to the exterior world. Thinkers in various fields, from philosophy to politics, agree that the definition is:

> ▷ A *concept* has
>   1. explicit and fixed content,

---

* 'Every word is a prejudice.' Friedrich Nietzsche (1844–1900), German philosopher.

** Ralph Waldo Emerson (1803–1882), US-American essayist and philosopher.

*** Insofar as one can say that mathematics is based on the concepts of 'set' and 'relation', which are based on experience, one can say that mathematics explores a section of reality, and that its concepts are *derived* from experience. This and similar views of mathematics are called *platonism*. More concretely, platonism is the view that the concepts of mathematics exist *independently* of people, and that they are discovered, and not created, by mathematicians.

In short, since mathematics makes use of the brain, which is a physical system, actually *mathematics is applied physics*.

However, we will discover that the concept of 'set' does *not* apply to nature; this changes the discussion in completely.

2. explicit and fixed limits,

3. explicit and fixed domain of application.

The inability to state these properties or to keep them fixed is often the easiest way to distinguish *crackpots* from more reliable thinkers. Unclearly defined terms, which thus do not qualify as concepts, regularly appear in myths, e.g. 'dragon' or 'sphinx', or in ideologies, e.g. 'worker' or 'soul'. Even physics is not immune. For example, we will discover
later that neither 'universe' nor 'creation' are concepts. Are you able to argue the case?

But the three defining properties of any concept are interesting in their own right. Explicit content means that concepts are built one onto another. In particular, the most fundamental concepts appear to be those that have no parts and no external relations,
but only internal ones. Can you think of one? Only the last part of this walk will uncover the final word on the topic.

The requirements of explicit limits and explicit contents also imply that all concepts describing nature are *sets*, since sets obey the same requirements. In addition, explicit domains of application imply that all concepts also are *relations*.* Since mathematics is based on the concepts of 'set' and of 'relation', one follows directly that mathematics can provide the *form* for any concept, especially whenever high precision is required, as in the study of motion. Obviously, the *content* of the description is only provided by the study of nature itself; only then do concepts become useful.

In the case of physics, the search for sufficiently precise concepts can be seen as the single theme structuring the long history of the field. Regularly, new concepts have been proposed, explored in all their properties, and tested. Finally, concepts are rejected or adopted, in the same way that children reject or adopt a new toy. Children do this unconsciously, scientists do it consciously, using language.** For this reason, concepts are universally intelligible.

Note that the concept '*concept*' itself is not definable independently of experience; a concept is something that helps us to act and react to the world in which we live. Moreover, concepts do not live in a world separate from the physical one: every concept requires memory from its user, since the user has to remember the way in which it was formed; therefore every concept needs a material support for its use and application. Thus all thinking and all science is fundamentally based on experience.

In conclusion, all concepts are based on the idea that nature is made of related parts. This idea leads to complementing couples such as 'noun–verb' in linguistics, 'set–relation' or 'definition–theorem' in mathematics, and 'aspect of nature–pattern of nature' in physics. These couples constantly guide human thinking, from childhood onwards, as developmental psychology can testify.

---

* We see that every physical concept is an example of a (mathematical) *category*, i.e., a combination of objects and mappings. For more details about categories, with a precise definition of the term, see page 199.
** Concepts formed unconsciously in our early youth are the most difficult to define precisely, i.e., with language. Some who were unable to define them, such as the Prussian philosopher Immanuel Kant (1724 –1804) used to call them 'a priori' concepts (such as 'space' and 'time') to contrast them with the more clearly defined 'a posteriori' concepts. Today, this distinction has been shown to be unfounded both by the study of child psychology (see the footnote on page 177) and by physics itself, so that these qualifiers are therefore not used in our walk.
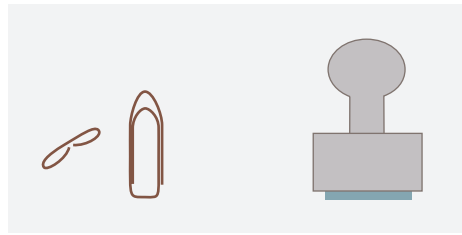
FIGURE 114 Devices for the definition of sets (left)
and of relations (right).

What are sets? What are relations?

> Alles, was wir sehen, könnte auch anders sein.
> Alles, was wir überhaupt beschreiben können,
> könnte auch anders sein. Es gibt keine Ordnung
> der Dinge a priori.*
>
> Ludwig Wittgenstein, *Tractatus*, 5.634

Defining sets and defining relations are the two fundamental acts of our thinking. This can be seen most clearly in any book about mathematics; such a book is usually divided into paragraphs labelled 'definition', 'theorem', 'lemma' and 'corollary'. The first type of paragraph defines concepts, i.e., defines sets, and the other three types of paragraphs express relations, i.e., connections between these sets. *Mathematics* is thus the exploration of the possible symbolic concepts and their relations. Mathematics is the science of symbolic necessities.

Sets and relations are tools of classification; that is why they are also the tools of any bureaucrat. (See Figure 114.) This class of humans is characterized by heavy use of paper clips, files, metal closets, archives – which all define various types of sets – and by the extensive use of numbers, such as reference numbers, customer numbers, passport numbers, account numbers, law article numbers – which define various types of relations between the items, i.e., between the elements of the sets.

Both the concepts of set and of relation express, in different ways, the fact that nature can be *described*, i.e., that it can be classified into parts that form a whole. The act of grouping together aspects of experience, i.e., the act of classifying them, is expressed in formal language by saying that a set is defined. In other words, a *set* is a collection of *elements* of our thinking. Every set distinguishes the elements from each other and from the set itself. This definition of 'set' is called the *naive* definition. For physics, the definition is sufficient, but you won't find many who will admit this. In fact, mathematicians have refined the definition of the concept 'set' several times, because the naive definition does not work well for infinite sets. A famous example is the story about sets which do not contain themselves. Obviously, any set is of two sorts: either it contains itself or it does not. If we take the set of all sets that do *not* contain themselves, to which sort does it belong?

To avoid problems with the concept of 'set', mathematics requires a precise definition. The first such definition was given by the German mathematician Ernst Zermelo (b. 1871 Berlin, d. 1951 Freiburg i.B.) and the German–Israeli mathematician Adolf/Abraham

---

* 'Everything we see could also be otherwise. Everything we describe at all could also be otherwise. There is no order of things a priori.'

TABLE 20 The defining properties of a set – the ZFC axioms

THE AXIOMS OF ZERMELO–FRAENKEL–C SET THEORY
– Two sets are equal if and only if they have the same elements. (Axiom of extensionality)
– The empty set is a set. (Axiom of the null set)
– If x and y are sets, then the unordered pair $\{x, y\}$ is a set. (Axiom of unordered pairs)
– If x is a set of sets, the union of all its members is a set. (Union or sum set axiom)
– The entity $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, ...\}$ is a set [a] – in other words, infinite collections, such as the natural numbers, are sets. (Axiom of infinity)
– An entity defined by all elements having a given property is a set, provided this property is reasonable; some important technicalities defining 'reasonable' are necessary. (Axiom of separation)
– If the domain of a function is a set, so is its range. (Axiom of replacement)
– The entity y of all subsets of x is also a set, called the power set. (Axiom of the power set)
– A set is not an element of itself – plus some technicalities. (Axiom of regularity)
– The product of a family of non-empty sets is non-empty. Equivalently, picking elements from a list of sets allows one to construct a new set – plus technicalities. (Axiom of choice)

a. The more common formulation (though equivalent to the above) is: The entity $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, ...\}$ is a set.

Fraenkel (b. 1891 München, d. 1965 Jerusalem). Later, the so-called *axiom of choice* was added, in order to make it possible to manipulate a wider class of infinite sets. The result of these efforts is called the *ZFC* definition.[*] From this basic definition we can construct all mathematical concepts used in physics. From a practical point of view, it is sufficient to keep in mind that for the whole of physics, the naive definition of a set is equivalent to the precise ZFC definition, actually even to the simpler ZF definition. Subtleties appear only for some special types of infinite sets, but these are not used in physics. In short, from the basic, naive set definition we can construct all concepts used in physics.

The naive set definition is far from boring. To satisfy two people when dividing a cake, we follow the rule: I cut, you choose. The method has two properties: it is *just*, as everybody thinks that they have the share that they deserve, and it is *fully satisfying*, as everybody has the feeling that they have at least as much as the other. What rule is needed for three people? And for four?

Apart from defining sets, every child and every brain creates links between the different aspects of experience. For example, when it hears a voice, it automatically makes the connection that a human is present. In formal language, connections of this type are

---

[*] A global overview of axiomatic set theory is given by PAUL J. COHEN & REUBEN HERSCH, Non-Cantorian set theory, *Scientific American* 217, pp. 104–116, 1967. Those were the times when *Scientific American* was a quality magazine.

For a good introduction to the axiom of choice, see the www.math.vanderbilt.edu/~schectex/ccc/choice.html website.

Other types of entities, more general than standard sets, obeying other properties, can also be defined, and are also subject of (comparatively little) mathematical research. To find an example, see the section on cardinals later on. Such more general entities are called classes whenever they contain at least one set. Can you give an example? In the final part of our mountain ascent we will meet physical concepts that are described neither by sets nor by classes, containing no set at all. That is where the real fun starts.

called *relations*. Relations connect and differentiate elements along other lines than sets: the two form a complementing couple. Defining a set unifies many objects and at the same time divides them into two: those belonging to the set and those that do not; defining a (binary) relation unifies elements two by two and divides them into many, namely into the many couples it defines.

Sets and relations are closely interrelated concepts. Indeed, one can define (mathematical) relations with the help of sets. A *(binary) relation* between two sets X and Y is a subset of the product set, where the *product set* or *Cartesian product $X \times Y$* is the set of all ordered pairs (x, y) with $x \in X$ and $y \in Y$. An *ordered pair* (x, y) can easily be defined with the help of sets. Can you find out how? For example, in the case of the relation 'is wife of', the set X is the set of all women and the set Y that of all men; the relation is given by the list all the appropriate ordered pairs, which is much smaller than the product set, i.e., the set of all possible woman–man combinations.

It should be noted that the definition of relation just given is not really complete, since every construction of the concept 'set' already contains certain relations, such as the relation 'is element of'. It does not seem to be possible to reduce either one of the concepts 'set' or 'relation' completely to the other one. This situation is reflected in the physical cases of sets and relations, such as space (as a set of points) and distance, which also seem impossible to separate completely from each other. In other words, even though mathematics does not pertain to nature, its two basic concepts, sets and relations, are taken from nature. In addition, the two concepts, like those of space-time and particles, are each defined with the other.

Infinity

Mathematicians soon discovered that the concept of 'set' is only useful if one can also call collections such as $\{0, 1, 2, 3...\}$, i.e., of the number 0 and all its successors, a 'set'. To achieve this, one property in the Zermelo–Fraenkel list defining the term 'set' explicitly specifies that this collection can be called a set. (In fact, also the axiom of replacement states that sets may be infinite.) Infinity is thus put into mathematics and into the tools of our thought right at the very beginning, in the definition of the term 'set'. When describing nature, with or without mathematics, we should never forget this fact. A few additional points about infinity should be of general knowledge to any expert on motion.

Only *sets* can be infinite. And sets have parts, namely their elements. When a thing or a concept is called 'infinite' one can *always* ask and specify what its parts are: for space the parts are the points, for time the instants, for the set of integers the integers, etc. An indivisible or a finitely divisible entity cannot be called infinite.*

A set is infinite if there is a function from it into itself that is *injective* (i.e., different elements map to different results) but not *onto* (i.e., some elements do not appear as images of the map); e.g. the map $n \mapsto 2n$ shows that the set of integers is infinite. Infinity also can be checked in another way: a set is infinite if it remains so also after removing one element, even repeatedly. We just need to remember that the empty set is *finite*.

---

* Therefore, most gods, being concepts and thus sets, are either finite or, in the case where they are infinite, they are divisible. It seems that only polytheistic world views are not disturbed by this conclusion.

There are *many types* of infinities, all of different sizes.* This important result was discovered by the Danish-Russian-German mathematician Georg Cantor (1845–1918). He showed that from the countable set of natural numbers one can construct other infinite sets which are not countable. He did this by showing that the *power set* $P(\omega)$, namely the set of all subsets, of a countably infinite set is infinite, but *not* countably infinite. Sloppily speaking, the power set is 'more infinite' than the original set. The real numbers $\mathbb{R}$, to be defined shortly, are an example of an uncountably infinite set; there are many more of them than there are natural numbers. (Can you show this?) However, *any* type of infinite set contains at least one subset which is countably infinite.

Even for an infinite set one can define size as the number of its elements. Cantor called this the *cardinality* of a set. The cardinality of a finite set is simply given by the number of its elements. The cardinality of a power set is 2 exponentiated by the cardinality of the set. The cardinality of the set of integers is called $\aleph_0$, pronounced 'aleph zero', after the first letter of the Hebrew alphabet. The smallest *uncountable* cardinal is called $\aleph_1$. The next cardinal is called $\aleph_2$ etc. A whole branch of mathematics is concerned with the manipulation of these infinite 'numbers'; addition, multiplication, exponentiation are easily defined. For some of them, even logarithms and other functions make sense.**

The cardinals defined in this way, including $\aleph_n$, $\aleph_\omega$, $\aleph_{\aleph_\aleph}$ are called accessible, because since Cantor, people have defined even larger types of infinities, called *inaccessible*. These numbers (inaccessible cardinals, measurable cardinals, supercompact cardinals, etc.) need additional set axioms, extending the ZFC system. Like the ordinals and the cardinals, they form examples of what are called *transfinite* numbers.

The real numbers have the cardinality of the power set of the integers, namely $2^{\aleph_0}$. Can you show this? The result leads to the famous question: Is $\aleph_1 = 2^{\aleph_0}$ or not? The statement that this be so is called the *continuum hypothesis* and was unproven for several generations. The surprising answer came in 1963: the usual definition of the concept of set is not specific enough to fix the answer. By specifying the concept of set in more detail, with additional axioms – remember that axioms are defining properties – you can make the continuum hypothesis come out either right or wrong, as you prefer.

Another result of research into transfinites is important: for every definition of a type of infinite cardinal, it seems to be possible to find a larger one. In everyday life, the idea of infinity is often used to stop discussions about size: 'My big brother is stronger than yours.' 'But mine is infinitely stronger than yours!' Mathematics has shown that questions on size do continue afterwards: 'The strength of my brother is the power set of that of yours!' Rucker reports that mathematicians conjecture that there is no possible nor any conceivable end to these discussions.

For physicists, a simple question appears directly. Do infinite quantities exist in nature? Or better, is it necessary to use infinite quantities to describe nature? You might want to clarify your own opinion on the issue. It will be settled during the rest of our adventure.

---

* In fact, there is such a huge number of types of infinities that none of these infinities itself actually describes this number. Technically speaking, there are as many infinities as there are ordinals.

** Many results are summarized in the excellent and delightful paperback by RUDY RUCKER, *Infinity and the Mind – the Science and Philosophy of the Infinite*, Bantam, Toronto, 1983.

FUNCTIONS AND STRUCTURES

Which relations are useful to describe patterns in nature? A typical example is 'larger stones are heavier'. Such a relation is of a specific type: it relates one specific value of an observable 'volume' to one specific value of the observable 'weight'. Such a one-to-one relation is called a *(mathematical) function* or *mapping*. Functions are the most specific types of relations; thus they convey a maximum of information. In the same way as numbers are used for observables, functions allow easy and precise communication of relations between observations. All physical rules and 'laws' are therefore expressed with the help of functions and, since physical 'laws' are about measurements, functions of numbers are their main building blocks.

A *function* f, or *mapping*, is a thus binary relation, i.e., a set $f = \{(x, y)\}$ of ordered pairs, where for every value of the first element x, called the *argument*, there is only *one* pair (x, y). The second element y is called the *value* of the function at the argument x. The set X of all arguments x is called the *domain of definition* and the set Y of all second arguments y is called the *range* of the function. Instead of $f = \{(x, y)\}$ one writes

$$f : X \to Y \quad \text{and} \quad f : x \mapsto y \quad \text{or} \quad y = f(x), \qquad (94)$$

where the type of arrow – with initial bar or not – shows whether we are speaking about sets or about elements.

We note that it is also possible to use the couple 'set' and 'mapping' to define all mathematical concepts; in this case a relation is defined with the help of mappings. A modern school of mathematical thought formalized this approach by the use of (mathematical) *categories*, a concept that includes both sets and mappings on an equal footing in its definition.*

To think and talk more clearly about nature, we need to define more specialized concepts than sets, relations and functions, because these basic terms are too general. The most important concepts derived from them are operations, algebraic structures and numbers.

A *(binary) operation* is a function that maps the Cartesian product of two copies of a set X into itself. In other words, an operation w takes an ordered couple of arguments $x \in X$ and assigns to it a value $y \in X$:

$$w : X \times X \to X \quad \text{and} \quad w : (x, x) \mapsto y . \qquad (95)$$

Challenge 234 s Is division of numbers an operation in the sense just defined?

Now we are ready to define the first of three basic concepts of mathematics. An *algebraic structure*, also called an *algebraic system*, is (in the most restricted sense) a set

---

* A *category* is defined as a collection of objects and a collection of 'morphisms', or mappings. Morphisms can be composed; the composition is associative and there is an identity morphism. The strange world of category theory, sometimes called the abstraction of all abstractions, is presented in WILLIAM LAWVERE & STEPHEN H. SCHANUEL, *Conceptual Mathematics: a First Introduction to Categories*, Cambridge University Press, 1997.

Note that every category contains a set; since it is unclear whether nature contains sets, as we will discuss on page 234, it is questionable whether categories will be useful in the unification of physics, despite their intense and abstract charm.

together with certain operations. The most important algebraic structures appearing in physics are groups, vector spaces, and algebras.

In addition to algebraic structures, mathematics is based on *order structures* and on *topological structures*. Order structures are building blocks of numbers and necessary to define comparisons of any sort. Topological structures are built, via subsets, on the concept of neighbourhood. They are necessary to define continuity, limits, dimensionality,
topological spaces and manifolds.

Obviously, most mathematical structures are combinations of various examples of these three basic structure types. For example, the *system* of real numbers is given by the *set* of real numbers with the *operations* of addition and multiplication, the *order relation* 'is larger than' and a *continuity* property. They are thus built by combining an algebraic
structure, an order structure and a topological structure. Let us delve a bit into the details.

Numbers

> Which numbers are multiplied by six when their last digit is taken away and transferred to the front?

> Never ask a man how many languages he has learned, how many countries he has seen, how much money he has accumulated, or how many women he has loved. If he can give you a precise answer, it means it was not enough.
>
> Jacques Mayol

Numbers are the oldest mathematical concept and are found in all cultures. The notion of number, in Greek ἀριθμός, has been changed several times. Each time the aim was to include wider classes of objects, but always retaining the general idea that numbers are entities that can be added, subtracted, multiplied and divided.

The modern way to write numbers, as e.g. in $12\,345\,679 \cdot 54 = 666\,666\,666$, is essential for science.* It can be argued that the lack of a good system for writing down and for calculating with numbers delayed the progress of science by several centuries. (By the way, the same can be said for the affordable mass reproduction of written texts.)

The simplest numbers, $0, 1, 2, 3, 4, ...$, are usually seen as being taken directly from experience. However, they can also be constructed from the notions of 'relation' and 'set'.
One of the many possible ways to do this (can you find another?) is by identifying a natural number with the set of its predecessors. With the relation 'successor of', abbreviated S, this definition can be written as

$$0 := \emptyset \quad , \quad 1 := S\,0 = \{0\} = \{\emptyset\} \quad ,$$
$$2 := S\,1 = \{0, 1\} = \{\emptyset, \{\emptyset\}\} \quad \text{and} \quad n + 1 := S\,n = \{0, ..., n\} \ . \tag{96}$$

This set, together with the binary operations 'addition' and 'multiplication,' constitutes the algebraic system $N = (N, +, \cdot, 1)$ of the *natural numbers*. For all number systems

---

* However, there is no need for written numbers for doing mathematics, as shown by MARCIA ASCHER, *Ethnomathematics – A Multicultural View of Mathematical Ideas*, Brooks/Cole, 1991.

the algebraic system and the set are often sloppily designated by the same symbol. The algebraic system N is what mathematician call a semi-ring. (Some authors prefer not to count the number zero as a natural number.) Natural numbers are fairly useful.

TABLE 21 Some large numbers.

NUMBER	EXAMPLE IN NATURE
Around us	
1	number of angels that can be in one place at the same time, following Thomas Aquinas Ref. 195
8	number of times a newspaper can be folded in alternate perpendicular directions
12	largest number of times a paper strip has been folded in the same direction Ref. 196
20	number of digits in precision measurements that will probably never be achieved
21, 34, 55, 89	petals of common types of daisy and sunflower Ref. 197
57	faces of a diamond with brilliant cut
2000	stars visible in the night sky
15 000	average number of objects in a European household
10^5	leaves of a tree (10 m beech)
6 to 7 $\cdot 10^9$	humans in the year 2000
10^{17}	ants in the world
$c.\ 10^{20}$	number of snowflakes falling on the Earth per year
$c.\ 10^{24}$	grains of sand in the Sahara desert
10^{22}	stars in the universe
10^{25}	cells on Earth
$1.1 \cdot 10^{50}$	atoms making up the Earth (6370^3 km$^3 \cdot 4 \cdot 3.14/3 \cdot 5500$ kg/m$^3 \cdot 30$ mol/kg $\cdot$ $6 \cdot 10^{23}$ /mol)
10^{81}	atoms in the visible universe
10^{90}	photons in the visible universe
10^{169}	number of atoms fitting in the visible universe
10^{244}	number of space-time points inside the visible universe
Information	
51	record number of languages spoken by one person
$c.\ 5000$	words spoken on an average day by a man
$c.\ 7000$	words spoken on an average day by a woman
$c.\ 2\,000\,000$	number of scientists on Earth around the year 2000
$3 \cdot 10^8$	words spoken during a lifetime (2/3 time awake, 30 words per minute)
$4 \cdot 10^9$	pulses exchanged between both brain halves every second
10^9	words heard and read during a lifetime
10^{17}	image pixels seen in a lifetime ($3 \cdot 10^9$ s $\cdot$ (1/15 ms) $\cdot$ 2/3 (awake) $\cdot 10^6$ (nerves to the brain) Ref. 198
10^{19}	bits of information processed in a lifetime (the above times 32)

Number	Example in nature
$c.\ 5 \cdot 10^{12}$	printed words available in (different) books around the world ($c.\ 100 \cdot 10^6$ books consisting of 50 000 words)
$2^{10} \cdot 3^7 \cdot 8! \cdot 12!$ $= 4.3 \cdot 10^{19}$	possible positions of the $3 \times 3 \times 3$ Rubik's Cube Ref. 199
$5.8 \cdot 10^{78}$	possible positions of the $4 \times 4 \times 4$ Rubik-like cube
$5.6 \cdot 10^{117}$	possible positions of the $5 \times 5 \times 5$ Rubik-like cube
$c.\ 10^{200}$	possible games of chess
$c.\ 10^{800}$	possible games of go
$c.\ 10^{10^7}$	possible states in a personal computer
Parts of us	
600	numbers of muscles in the human body, of which about half are in the face
$150\,000 \pm 50\,000$	hairs on a healthy head
900 000	neurons in the brain of a grasshopper
$126 \cdot 10^6$	light sensitive cells per retina (120 million rods and 6 million cones)
10^{10} to 10^{11}	neurons in the human brain
$> 10^{16}$	memory bits in the human brain
$500 \cdot 10^6$	blinks of the eye during a lifetime (about once every four seconds when awake)
$300 \cdot 10^6$	breaths taken during human life
$3 \cdot 10^9$	heart beats during a human life
$3 \cdot 10^9$	letters (base pairs) in haploid human DNA
$6.1 \cdot 10^9$	bits in a compact disc
$1 \cdot 10^{11}$	humans who have ever lived
$10^{15 \pm 1}$	cells in the human body
$10^{16 \pm 1}$	bacteria carried in the human body

The system of *integers* $Z = (\ldots, -2, -1, 0, 1, 2, \ldots, +, \cdot, 0, 1)$ is the minimal ring that is an extension of the natural numbers. The system of *rational numbers* $Q = (Q, +, \cdot, 0, 1)$ is the minimal field that is an extension of the ring of the integers. (The terms 'ring' and 'field' are defined in all details in the next volume.) The system of *real numbers* $R = (R, +, \cdot, 0, 1, >)$ is the minimal extension of the rationals that is continuous and totally ordered. (For the definition of continuity, see volume IV, and volume V, .) Equivalently, the reals are the minimal extension of the rationals forming a complete, totally strictly-Archimedean ordered field. This is the historical construction – or definition – of the integer, rational and real numbers from the natural numbers. However, it is not the only one construction possible. The most beautiful definition of all these types of numbers is the one discovered in 1969 by John Conway, and popularized by him, Donald Knuth and Martin Kruskal.

⊳ A *number* is a sequence of bits.

The two bits are usually called 'up' and 'down'. Examples of numbers and the way to write
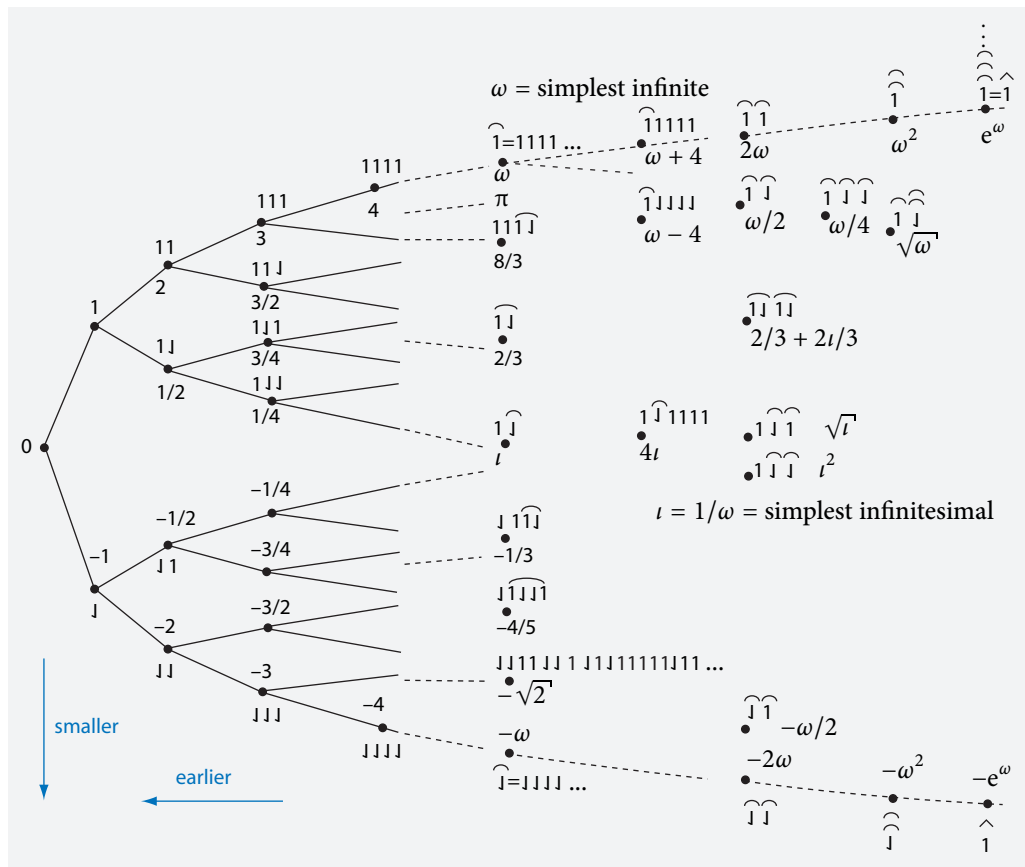
FIGURE 115 The surreal numbers in conventional and in bit notation.

them are given in Figure 115.

The empty sequence is the number zero. A finite sequence of n ups is the integer number n, and a finite sequence of n downs is the integer $-n$. Finite sequences of mixed ups and downs give the *dyadic rational numbers*. Examples are 1, 2, 3, −7, 19/4, 37/256, etc. They all have denominators with a power of 2. The other *rational numbers* are those that end in an infinitely repeating string of ups and downs, such as the *reals*, the *infinitesimals* and simple infinite numbers. Longer countably infinite series give even more crazy numbers.

The complete class of numbers that is defined by a sequence of bits is called the class of *surreal* numbers.[*]

There is a second way to write surreal numbers. The first is the just mentioned sequence of bits. But in order to define addition and multiplication, another notation is usually used, deduced from Figure 115. A surreal α is defined as the earliest number of

---

[*] The surreal numbers do *not* form a set since they contain all *ordinal numbers*, which themselves do not form a set, even though they of course *contain* sets. In short, ordinals and surreals are classes which are *larger* than sets.

all those between two series of earlier surreals, the left and the right series:

$$\alpha = \{a, b, c, ...|A, B, C, ...\} \quad \text{with} \quad a, b, c, < \alpha < A, B, C \ . \tag{97}$$

For example, we have

$$\{0|\} = 1 \ , \ \{0, 1\, |\} = 2 \ , \ \{|0\} = -1 \ , \ \{|-1, 0\} = -2 \ , \ \{0\,|1\} = 1/2 \ ,$$
$$\{0\,|1/2, 1/4\} = 1 \ , \ \{0, 1, 3/2, 25/16\,|41/16, 13/8, 7/4, 2\} = 1 + 37/64 \ , \tag{98}$$

showing that the finite surreals are the *dyadic numbers* $m/2^n$ (n and m being integers). Given two surreals $\alpha = \{..., a, ...|..., A, ...\}$ with $a < \alpha < A$ and $\beta = \{..., b, ...|..., B, ...\}$ with $b < \beta < B$, addition is defined recursively, using earlier, already defined numbers, as

$$\alpha + \beta = \{..., a + \beta, ..., \alpha + b, ...|..., A + \beta, ..., \alpha + B, ...\} \ . \tag{99}$$

This definition is used simply because it gives the same results as usual addition for integers and reals. Can you confirm this? By the way, addition is not always commutative. Are you able to find the exceptions, and to find the definition for subtraction? Multiplication is also defined recursively, namely by the expression

$$\alpha\beta = \{..., a\beta + \alpha b - ab, ..., A\beta + \alpha B - AB, ...|$$
$$..., a\beta + \alpha B - aB, ..., A\beta + \alpha b - Ab, ...\} \ . \tag{100}$$

These definitions allow one to write $\iota = 1/\omega$, and to talk about numbers such as $\sqrt{\omega}$, the square root of infinity, about $\omega + 4$, $\omega - 1$, 2ω, e^ω and about other strange numbers shown in Figure 115. However, the surreal numbers are not commonly used. More common is one of their subsets.

The *real numbers* are those surreals whose decimal expansion is not larger than infinity and in addition, equate numbers such as 0.999999... and 1.000000..., as well as all similar cases. In other words, the surreals distinguish the number 0.999999... from the number 1, whereas the reals do not. Indeed, between these two surreal numbers there are infinitely many other surreals. Can you name a few?

Reals are more useful for describing nature than surreals, first because they form a *set* – which the surreals do not – and secondly because they allow the definition of *integration*. Other numbers defined with the help of reals, e.g. the complex numbers $\mathbb{C}$, the quaternions $\mathbb{H}$ and a few more elaborate number systems, are presented in the next volume.

To conclude, in physics it is usual to call *numbers* the elements of any set that is a semi-ring (e.g. $\mathbb{N}$), a ring (e.g. $\mathbb{Z}$) or a field ($\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$). Since numbers allow one to compare magnitudes and thus to *measure*, these numbers play a central role in the description of observations.

“ A series of equal balls is packed in such a way
that the area of needed wrapping paper is
minimal. For small numbers of balls the linear
package, with all balls in one row, is the most
efficient. For which number of balls is the linear
package no longer a minimum?

Why use mathematics?

“ Die Forderung der Möglichkeit der einfachen
Zeichen ist die Forderung der Bestimmtheit des
Sinnes.*
Ludwig Wittgenstein, *Tractatus*, 3.23 ”

Several well-known physicists have repeatedly asked why mathematics is so important.
For example, Niels Bohr is quoted as having said: 'We do not know why the language
of mathematics has been so effective in formulating those laws in their most succinct
form.' Eugene Wigner wrote an often cited paper entitled *The unreasonable effectiveness
of mathematics.* At the start of science, many centuries earlier, Pythagoras and his con-
temporaries were so overwhelmed by the usefulness of numbers in describing nature,
that Pythagoras was able to organize a sect based on this connection. The members of
the inner circle of this sect were called 'learned people,' in Greek '*mathematicians*', from
the Greek μάθημα 'teaching'. This sect title then became the name of the modern profes-
sion.

These men forgot that numbers, as well as a large part of mathematics, are concepts
developed precisely with the aim of describing nature. Numbers and mathematical con-
cepts were developed right from the start to provide as succinct a description as possible.
That is one consequence of mathematics being the science of symbolic necessities.

Perhaps we are being too dismissive. Perhaps these thinkers mainly wanted to express
their feeling of wonder when experiencing that language works, that thinking and our
brain works, and that life and nature are so beautiful. This would put the title question
nearer to the well-known statement by Albert Einstein: 'The most incomprehensible fact
about the universe is that it is comprehensible.' Comprehension is another word for de-
scription, i.e., for classification. Obviously, any separable system is comprehensible, and
there is nothing strange about it. But is the universe separable? As long as is it described
as being made of particles and vacuum, this is the case.

We will find in the last part of this adventure that the basic assumption made at our
start is built on sand. The assumption that observations in nature can be *counted*, and
thus that nature is separable, is an approximation. The quoted 'incomprehensibility' be-
comes amazement at the precision of this approximation. Nevertheless, Pythagoras' sect,
which was based on the thought that 'everything in nature is numbers', was wrong. Like
so many beliefs, observation will show that it was wrong.

“ Die Physik ist für Physiker viel zu schwer.**
David Hilbert ”

---

* 'The requirement that simple signs be possible is the requirement that sense be determinate.'
** 'Physics is much too difficult for physicists.'

Is mathematics a language?

> Die Sätze der Mathematik sind Gleichungen,
> also Scheinsätze. Der Satz der Mathematik
> drückt keinen Gedanken aus.*
>   Ludwig Wittgenstein, *Tractatus*, 6.2, 6.21

Surely, mathematics is a *vocabulary* that helps us to talk with precision. Mathematics can be seen as the exploration of *all* possible concepts that can be constructed from the two fundamental bricks 'set' and 'relation' (or some alternative, but equivalent pair). *Mathematics* is the science of symbolic necessities. Rephrased again, mathematics is the exploration of all possible types of classifications. This explains its usefulness in all situations where complex, yet precise classifications of observations are necessary, such as in physics.

However, mathematics cannot express everything that humans want to communicate, such as wishes, ideas or feelings. Just try to express the fun of swimming using mathematics. Indeed, *mathematics is the science of symbolic necessities*; thus mathematics is not a language, nor does it contain one. Mathematical concepts, being based on *abstract* sets and relations, do not pertain to nature. Despite its beauty, mathematics does not allow us to talk about nature or the observation of motion. Mathematics does not tell *what* to say about nature; it does tell us *how* to say it.

In his famous 1900 lecture in Paris, the German mathematician David Hilbert** gave a list of 23 great challenges facing mathematics. The sixth of Hilbert's problems was to find a mathematical treatment of the axioms of physics. Our adventure so far has shown that physics started with a *circular definition* that has not yet been eliminated after 2500 years of investigations: space-time is defined with the help of objects and objects are defined

with the help of space and time. Being based on a circular definition, physics is thus *not* modelled after mathematics, even if many physicists and mathematicians, including
Hilbert, would like it to be so. Physicists must live with logical problems and must walk on unsure ground in order to achieve progress. In fact, they have done so for 2500 years. If physics were an axiomatic system, it would not contain circular definitions; on the other hand, it would also cease to be a language and would cease to describe nature. We
will return to this issue in the last part of our adventure.

---

* 'The propositions of mathematics are equations, and therefore pseudo-propositions. A proposition of mathematics does not express a thought.'
** David Hilbert (1862 Königsberg–1943 Göttingen) was professor of mathematics in Göttingen and the greatest mathematician of his time. He was a central figure to many parts of mathematics, and also played an important role both in the birth of general relativity and of quantum theory. His textbooks are still in print. His famous personal credo was: 'Wir müssen wissen, wir werden wissen.' (We must know, we will know.) His famous Paris lecture is published e.g. in *Die Hilbertschen Probleme*, Akademische Verlagsgesellschaft Geest & Portig, 1983. The lecture galvanized all of mathematics. (Despite efforts and promises of similar fame, *nobody* in the world had a similar overview of mathematics that allowed him or her to repeat the feat in the year 2000.) In his last decade he suffered the persecution of the Nazi regime; the persecution eliminated Göttingen from the list of important science universities, without recovering its place up to this day.

Curiosities and fun challenges about mathematics

Challenge 240 s What is the largest number that can be written with four digits of 2 and no other sign? And with four 4s?

* *

Challenge 241 e Pythagorean triplets are integers that obey $a^2 + b^2 = c^2$. Give at least ten examples. Then show the following three properties: at least one number in a triplet is a multiple of 3; at least one number in a triplet is a multiple of 4; at least one number in a triplet is a multiple of 5.

* *

Challenge 242 s A mother is 21 years older than her child, and in 6 years the child will be 5 times younger than the mother. Where is the father? This is the *young mother puzzle*.

* *

Challenge 243 d The number $1/n$, when written in decimal notation, has a periodic sequence of digits. The period is at most $n-1$ digits long, as for $1/7 = 0.142857\,142857\,1428....$ Which other numbers $1/n$ have periods of length $n-1$?

* *

Challenge 244 s Felix Klein was a famous professor of mathematics at Göttingen University. There were two types of mathematicians in his department: those who did research on whatever they wanted and those for which Klein provided the topic of research. To which type did Klein belong?

Challenge 245 s Obviously, this is a variation of another famous puzzle. A barber shaves all those people who do not shave themselves. Does the barber shave himself?

* *

Challenge 246 s Everybody knows what a *magic square* is: a square array of numbers, in the simplest case from 1 to 9, that are distributed in such a way that the sum of all rows, columns (and possibly all diagonals) give the same sum. Can you write down the simplest $3 \times 3 \times 3$ *magic cube*?

* *

Ref. 205 The digits 0 to 9 are found on keyboards in two different ways. Calculators and keyboards have the 7 at the top left, whereas telephones and automatic teller machines have the digit 1 at the top left. The two standards, respectively by the International Standards Organization (ISO) and by the International Telecommunication Union (ITU, formerly CCITT), evolved separately and have never managed to merge.

* *

Leonhard Euler in his notebooks sometimes wrote down equations like

$$1 + 2^2 + 2^4 + 2^6 + 2^8 + ... = -\frac{1}{3} \ . \tag{101}$$
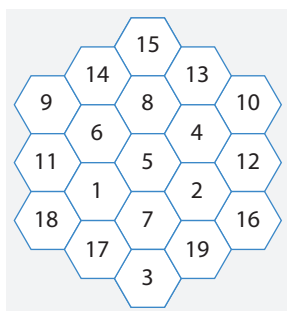
FIGURE 116 The only magic hexagon starting with the number 1 (up to reflections and rotations).
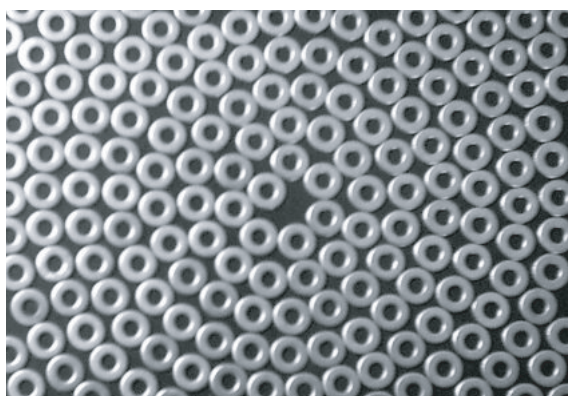
FIGURE 117 Fibonacci numbers and spirals from washers (© Donald Simanek).

Challenge 247 d   Can this make sense?

* *

In the history of recreational mathematics, several people have independently found the well-known magic hexagon shown in Figure 116. The first discoverer was, in 1887, Ernst von Hasselberg. The hexagon is called magic because all lines add up to the same number, 38. Hasselberg also proved the almost incredible result that no other magic
Challenge 248 d   hexagon exists. Can you confirm this?

* *

It is often said that numbers from the *Fibonacci series* 1, 1, 2, 3, 5, 8, 13, 21 etc., which
Ref. 197   appear so often in flowers, are due to some deep sense of beauty in nature. This is *not* the case, as Figure 117 shows. Mark a spot on a surface, and put washers around it in by hand in a spiral manner; you will find the same spirals that you find in many flowers, and thus, at their border, the same Fibonacci numbers.
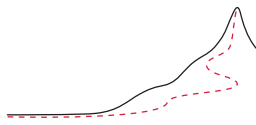
CHAPTER 9

CONCEPTS, LIES AND PATTERNS OF NATURE

> Die Grenzen meiner Sprache bedeuten die
> Grenzen meiner Welt.*
>   Ludwig Wittgenstein, *Tractatus*, 5.6

> Der Satz ist ein Bild der Wirklichkeit. Der Satz
> ist ein Modell der Wirklichkeit, so wie wir sie
> uns denken.**
>   Ludwig Wittgenstein, *Tractatus*, 4.01

I N contrast to mathematics, physics does aim at being a language. But
t is ambitious: it aims to express *everything*, with *complete* precision, and,
n particular, all examples and possibilities of change.*** Like any language, physics
consists of concepts and sentences. In order to be able to express everything, it must
aim to use few words for a lot of facts.**** Physicists are essentially *lazy* people: they
try to minimize the effort in everything they do. The concepts in use today have been
optimized by the combined effort of many people to be as practical, i.e., as powerful as
possible. A concept is called *powerful* when it allows one to express in a compact way
a large amount of information, meaning that it can rapidly convey a large number of
details about observations.

General statements about many examples of motion are called *rules* or *patterns*. In the
past, it was often said that 'laws govern nature', using an old and inappropriate ideology.
A physical 'law' is only a way of saying as much as possible with as few words as possible.

---

* 'The limits of my language are the limits of my world.'

** 'A proposition is a picture of reality. A proposition is a model of reality as we imagine it.'

*** All observations are about change. The various types of change are studied by the various sciences; they
are usually grouped in the three categories of *human sciences*, *formal sciences* and *natural sciences*. Among
the latter, the oldest are astronomy and metallurgy. Then, with the increase of curiosity in early antiquity,
came the natural science concerned with the topic of motion: *physics*. In the course of our walk it will
become clear that the unusual definition of physics as the study of change indeed covers the whole set of
topics studied in physics. In particular it includes the more common definition of physics as the study of
matter, its properties, its components and their interactions.

**** A particular, specific observation, i.e., a specific example of input shared by others, is called a *fact*, or
in other contexts, an *event*. A striking and regularly observed fact is called a *phenomenon*, and a general
observation made in many different situations is called a *(physical) principle*. (Often, when a concept is in-
troduced that is used with other meaning in other fields, in this walk it is preceded by the qualifier 'physical'
or 'mathematical' in parentheses.) Actions performed towards the aim of collecting observations are called
experiments. The concept of experiment became established in the sixteenth century; in the evolution of a
child, it can best be compared to that activity that has the same aim of collecting experiences: *play*.

When saying 'laws govern nature' we actually mean to say 'being lazy, we describe observations with patterns'. Laws are the epitome of laziness. Formulating laws is pure sloth. In fact, the correct expression is *patterns describe nature*.

Physicists have defined the laziness necessary for their field in much detail. In order to become a master of laziness, we need to distinguish lazy patterns from those which are not, such as lies, beliefs, statements that are not about observations, and statements that are not about motion. We do this below.

The principle of extreme laziness is the origin, among others, of the use of numbers in physics. Observables are often best described with the help of numbers, because numbers allow easy and precise communication and classification. Length, velocity, angles, temperature, voltage or field strength are of this type. The notion of 'number', used in every measurement, is constructed, often unconsciously, from the notions of 'set' and 'relation', as shown above. Apart from the notion of number, other concepts are regularly defined to allow fast and compact communication of the 'laws' of nature; all are 'abbreviation tools.' In this sense, the statement 'the level of the Kac–Moody algebra of the Lagrangian of the heterotic superstring model is equal to one' contains precise information, explainable to everybody; however, it would take dozens of pages to express it using only the terms 'set' and 'relation.' In short, the *precision* common in physics results from its *quest for laziness*.

> Es ist besser, daß die Leute nicht wissen, wie Gesetze und Wurst zustande kommen. Sonst könnten sie nachts nicht ruhig schlafen.*
>
> Otto von Bismarck

ARE PHYSICAL CONCEPTS DISCOVERED OR CREATED?

> Das logische Bild der Tatsachen ist der Gedanke.**
>
> Ludwig Wittgenstein, *Tractatus*, 3

The title question is often rephrased as: are physical concepts free of beliefs, taste or personal choices? The question has been discussed so much that it even appears in Hollywood films. We give a short summary that can help you to distinguish honest from dishonest teachers.

Creation of concepts, in contrast to their discovery, would imply free choice between many alternative possibilities. The chosen alternative would then be due to the beliefs or tastes used. In physics (in obvious contrast to other, more ideological fields of enquiry), we know that different physical descriptions of observations are either equivalent or, in the opposite case, imprecise or even wrong. A description of observations is thus essentially unique: any choices of concepts are only apparent. There is no real freedom in the definition of physical concepts. In this property, physics is in strong contrast to artistic activity.

If two different concepts can be used to describe the same aspect of observations, they must be equivalent, even if the relation that leads to the equivalence is not immediately clear. In fact, the requirement that people with different standpoints and observing the

---

* 'It is better that people do not know how laws and sausages are made. Otherwise they would not sleep well at night.' Otto von Bismarck (1815–1898), Prussian Chancellor.
** 'A logical picture of facts is a thought.'

same event deduce equivalent descriptions lies at the very basis of physics. It expresses the requirement that observations are observer independent. In short, the strong requirement of viewpoint independence makes the free choice of concepts a logical impossibility.

The conclusion that concepts describing observations are discovered rather than created is also reached independently in the field of linguistics by the above-mentioned research on semantic primitives,* in the field of psychology by the observations on the formation of the concepts in the development of young children, and in the field of ethology by the observations of animal development, especially in the case of mammals. In all three fields detailed observations have been made of how the interactions between an individual and its environment lead to concepts, of which the most basic ones, such as space, time, object or interaction, are common across the sexes, cultures, races and across many animal species populating the world. Curiosity and the way that nature works leads to the same concepts for all people and even the animals; the world offers only one possibility, without room for imagination. Imagining that physical concepts can be created at your leisure is a belief – or a useful exercise, but never successful.

Physical concepts are classifications of observations. The activity of classification itself follows the patterns of nature; it is a mechanical process that machines can also perform. This means that any distinction, i.e., any statement that A is different from B, is a theory-free statement. No belief system is necessary to distinguish different entities in nature. Cats and pigs can also do so. Physicists *can* be replaced by animals, even by machines. Our mountain ascent will repeatedly confirm this point.

As already mentioned, the most popular physical concepts allow us to describe observations as succinctly and as accurately as possible. They are formed with the aim of having the largest possible amount of understanding with the smallest possible amount of effort. Both Occam's razor – the requirement not to introduce unnecessary concepts – and the drive for unification automatically reduce the number and the type of concepts used in physics. In other words, the progress of physical science was and is based on a programme that reduces the possible choice of concepts as drastically as possible.

In summary, we found that physical concepts are the same for everybody and are free of beliefs and personal choices: they are first of all *boring*. Moreover, as they could stem from machines instead of people, they are *born of laziness*. Despite these human analogies – not meant to be taken too seriously – physical concepts are *not* created; they are discovered. If a teacher tells you the opposite, he is lying.

Having handled the case of physical concepts, let us now turn to physical statements. The situation is somewhat similar: physical statements must be lazy, arrogant and boring. Let us see why.

> “Wo der Glaube anfängt, hört die Wissenschaft auf.**
> Ernst Haeckel, *Natürliche Schöpfungsgeschichte*, 1879.

---

* Anna Wierzbicka concludes that her research clearly indicates that semantic primitives are *discovered*, in particular that they are deduced from the fundamentals of human experience, and not invented.
** Where belief starts, science ends.

TABLE 22 The 'scientific method'.

Normal description **Curiosity**	Lobbyist description **Scientific method**
1. look around a lot	1. interact with the world
2. don't believe anything told	2. forget unproven statements
3. choose something interesting and explore it yourself	3. observe
4. make up your own mind and describe precisely what you saw	4. use reason, build hypothesis
5. check if you can also describe similar situations in the same way	5. analyse hypothesis
6. increase the precision of observation until the checks either fail or are complete	6. perform experiments to check hypothesis
7. depending on the case, continue with step 4 or 1	7. ask authority for more money

HOW DO WE FIND PHYSICAL PATTERNS AND RULES?

> Grau, theurer Freund, ist alle Theorie,
> Und grün des Lebens goldner Baum.*
> > J.W. v. Goethe, *Faust*.

> Physics is usually presented as an objective
> science, but I notice that physics changes and
> the world stays the same, so there must be
> something subjective about physics.
> > Richard Bandler

Progressing through the study of motion reflects a young child's attitude towards life. The progress follows the simple programme on the left of Table 22.

Adult scientists do not have much more to add, except the more fashionable terms on the right, plus several specialized professions to make money from them. The experts of step 7 are variously called lobbyists or fund raisers; instead of calling this program 'curiosity', they call it the 'scientific method.' They mostly talk. Physics being the talk about motion,** and motion being a vast topic, many people specialize in this step.

The experts of step 6 are called *experimental physicists* or simply *experimentalists*, a term derived from the Latin 'experiri', meaning 'to try out'. Most of them are part of the category 'graduate students'. The experts of steps 5 and 4 are called *theoretical physicists*

---

* 'Grey, dear friend, is all theory, and green the golden tree of life.' Johann Wolfgang von Goethe (1749–1832), the influential German poet.
** Several sciences have the term 'talk' as part of their name, namely all those whose name finishes in '-logy', such as e.g. biology. The ending stems from ancient Greek and is deduced from λήγηιν meaning 'to say, to talk'. Physics as the science of motion could thus be called 'kinesiology' from κίνησις, meaning 'motion'; but for historical reasons this term has a different meaning, namely the study of human muscular activity. The term 'physics' is either derived from the Greek φύσικη (τέχνη is understood) meaning '(the art of) nature', or from the title of Aristotle' works τὰ φυσικά meaning 'natural things'. Both expressions are derived from φύσις, meaning 'nature'.

or simply *theoreticians*.* This is a rather modern term; for example, the first professors of theoretical physics were appointed around the start of the twentieth century. The term is derived from the Greek θεωρία meaning 'observation, contemplation'. Finally, there are the people who focus on steps 1 to 3, and who induce others to work on steps 4 to 6; they are called *geniuses*.

Obviously an important point is hidden in step 6: how do all these people know whether their checks fail? How do they recognize truth?

> All professions are conspiracies against laymen.
>
> George Bernard Shaw

What is a lie?

> Get your facts straight, and then you can distort them at your leisure.
>
> Mark Twain

> The pure truth is always a lie.
>
> Bert Hellinger

Lies are useful statements, as everybody learns during their youth. One reason that they are useful is because we can draw any imaginable conclusion from them. A well-known discussion between two Cambridge professors early in the twentieth century makes the point. McTaggart asked: 'If 2 + 2 = 5, how can you prove that I am the pope?' Godfrey Hardy: 'If 2 + 2 = 5, then 4 = 5; subtract 3; then 1 = 2; but McTaggart and the pope are two; therefore McTaggart and the pope are one.' As noted long ago, *ex falso quodlibet*; from what is wrong, anything imaginable can be deduced. It is true that in our mountain ascent we need to build on previously deduced results and that our trip could not be completed if we had a false statement somewhere in our chain of arguments. But lying is such an important activity that one should learn to perform it well.

Ref. 163   There are various stages in the art of lying. Many animals have been shown to deceive their kin. Children start lying just before their third birthday, by hiding experiences. Adults cheat on taxes. And many intellectuals or politicians even claim that truth does not exist. There are also people who have spent most of their research career on lies and lying.**

However, in most countries, everybody must know what 'truth' is, since in a law court for example, telling an untruth can lead to a prison sentence. The courts are full of experts in lie detection. If you lie in court, you better do it well; experience shows that you might get away with many criminal activities. In court, a *lie* is a statement that knowingly contrasts with observations.*** The truth of a statement is thus checked by observation.

---

* If you like theoretical physics, have a look at the refreshingly candid web page by Nobel Prize winner Gerard 't Hooft with the title *How to become a good theoretical physicist*. It can be found at www.phys.uu. nl/~thooft/theorist.html.

** A well-known example is Paul Ekman, whose fascinating website at www.paulekman.com tells how to spot lies from the behaviour of the person telling it.

*** Statements not yet checked are variously called *speculations*, *conjectures*, *hypotheses*, or – wrongly – simply *theses*. Statements that are in correspondence with observations are called *correct* or *true*; statements that contrast with observations are called *wrong* or *false*.

The check itself is sometimes called the *proof* of the statement. For law courts, as for physics, *truth* is thus the correspondence with facts, and *facts* are shared observations. A 'good' lie is thus a lie whose contrast with shared observations is hard to discover.

The first way of lying is to put an emphasis on the sharedness only. Populists and polemics do this regularly. ('Every foreigner is a danger for the values of our country.') Since almost any imaginable opinion, however weird, is held by some group – and thus shared – one can always claim it as true.* Unfortunately, it is no secret that ideas also get shared because they are fashionable, imposed or opposed to somebody who is generally disliked. Often a sibling in a family has this role – remember Cassandra.** For a good lie we thus need more than sharedness, more than *intersubjectivity* alone.

A good lie should be, like a true statement, really independent of the listener and the observer and, in particular, independent of their age, their sex, their education, their civilization or the group to which they belong. For example, it is especially hard – but not impossible – to lie with mathematics. The reason is that the basic concepts of mathematics, be they 'set', 'relation' or 'number', are taken from observation and are intersubjective, so that statements about them are easily checked. Usually, lies thus avoid mathematics.***

Secondly, a 'good' lie should avoid statements about observations and use *interpretations* instead. For example, some people like to talk about other universes, which implies talking about fantasies, not about observations. A good lie has to avoid, however, to fall in the opposite extreme, namely to make statements which are meaningless; the most destructive comment that can be made about a statement is the one used by the great Austrian physicist Wolfgang Pauli: 'That is not even wrong.'

Thirdly, a good lie doesn't care about observations, only about imagination. Only truth needs to be *empirical*, to distinguish it from *speculative* statements. If you want to lie 'well' even with empirical statements, you need to pay attention. There are two types of empirical statements: *specific* statements and *universal* statements. For example, 'On the 31st of August 1960 I saw a green swan swimming on the northern shore of the lake of Varese' is specific, whereas 'All ravens are black' is universal, since it contains the term 'all'. There is a well-known difference between the two, which is important for lying well: specific statements cannot be falsified, they are only verifiable, and universal statements

cannot be verified, they are only falsifiable. Why is this so?

Universal statements such as 'the speed of light is constant' cannot be tested for *all* possible cases. (Note that if they could, they would not be universal statements, but just a list of specific ones.) However, they can be reversed by a counter-example. Another example of the universal type is: 'Apples fall upwards.' Since it is falsified by an observation conducted by Newton several centuries ago, or by everyday experience, it qualifies

---

* The work of the French sociologist Gabriel Tarde (1843–1903), especially his concepts of *imitation* and group mind, already connects to this fact.

** The implications of birth order on creativity in science and on acceptance of new ideas has been studied in the fascinating book by Frank J. Sulloway, *Born to Rebel – Birth Order, Family Dynamics and Creative Lives*, Panthon Books, 1996. This exceptional book tells the result of a life-long study correlating the personal situations in the families of thousands of people and their receptivity to about twenty revolutions in the recent history. The book also includes a test in which the reader can deduce their own propensity to rebel, on a scale from 0 to 100 %. Darwin scores 96 % on this scale.

*** In mathematics, 'true' is usually specified as 'deducible' or 'provable'; this is in fact a special case of the usual definition of truth, namely 'correspondence with facts', if one remembers that mathematics studies the properties of classifications.

as an (easily detectable) lie. In general therefore, lying by stating the opposite of a theory is usually unsuccessful. If somebody insists on doing so, the lie becomes a *superstition*, a *belief*, a *prejudice* or a *doctrine*. These are the low points in the art of lying. A famous case of insistence on a lie is that of the colleagues of Galileo, who are said to have refused to look through his telescope to be convinced that Jupiter has moons, an observation that would have shaken their belief that everything turns around the Earth. Obviously these astronomers were amateurs in the art of lying. A good universal lie is one whose counter-example is not so easily spotted.

There should be no insistence on lies in physics. Unfortunately, classical physics is full of lies. We will dispel them during the rest of our walk.

Lying by giving specific instead of universal statements is much easier. ('I can't remember.') Even a specific statement such as 'yesterday the Moon was green, cubic and smelled of cheese' can never be completely falsified: there is no way to show with absolute certainty that this is wrong. The only thing that we can do is to check whether the statement is compatible with other observations, such as whether the different shape affected the tides as expected, whether the smell can be found in air collected that day, etc. A good specific lie is thus not in contrast with other observations.*

Incidentally, universal and specific statements are connected: the *opposite* of a universal statement is always a specific statement, and vice versa. For example, the opposite of the general statement 'apples fall upwards', namely 'some apples fall downwards', is specific. Similarly, the specific statement 'the Moon is made of green cheese' is in opposition to the universal statement 'the Moon is solid for millions of years and has almost no smell or atmosphere.'

In other words, law courts and philosophers disagree. Law courts have no problem with calling theories true, and specific statements lies. Many philosophers avoid this. For example, the statement 'ill-tempered gaseous vertebrates do not exist' is a statement of the universal type. If a universal statement is in agreement with observations, and if it is falsifiable, law courts call it *true*. The opposite, namely the statement: 'ill-tempered gaseous vertebrates do exist', is of the *specific* type, since it means 'Person X has observed an ill-tempered gaseous vertebrate in some place Y at some time Z'. To verify this, we need a record of the event. If such a record, for example a photographs or testimony does not exist, and if the statement *can* be falsified by other observations, law courts call the specific statement a *lie*. Even though these are the rules for everyday life and for the law, there is no agreement among philosophers and scientists that this is acceptable.

---

* It is often difficult or tedious to verify statements concerning the past, and the difficulty increases with the distance in time. That is why people can insist on the occurrence of events which are supposed to be exceptions to the patterns of nature ('miracles'). Since the advent of rapid means of communication these checks are becoming increasingly easy, and no miracles are left over. This can be seen in Lourdes in France, where even though today the number of visitors is much higher than in the past, no miracles have been seen in decades.

In fact, all modern 'miracles' are kept alive only by consciously eschewing checks, such as the supposed yearly liquefaction of blood in Napoli, the milk supposedly drunk by statues in temples, the supposed healers in television evangelism, etc. Most miracles only remain because many organizations make money out of the difficulty of falsifying specific statements. For example, when the British princess Diana died in a car crash in 1997, even though the events were investigated in extreme detail, the scandal press could go on almost without end about the 'mysteries' of the accident.

Why? Intellectuals are a careful lot, because many of them have lost their lives as a result of exposing lies too openly.

In short, specific lies, like all specific statements, can never be falsified with certainty. This is what makes them so popular. Children learn specific lies first. ('I haven't eaten the jam.') General lies, like all general statements, can always be corroborated by examples. This is the reason for the success of ideologies. But the criteria for recognizing lies, even general lies, have become so commonplace that beliefs and lies try to keep up with them. It became fashionable to use expressions such as 'scientific fact' – there are no non-scientific facts –, or 'scientifically proven' – observations cannot be proven otherwise – and similar empty phrases. These are not 'good' lies; whenever we encounter sentences beginning with 'science says ...' or 'science and religion do ...', replacing 'science' by 'knowledge' or 'experience' is an efficient way of checking whether such statements are to be taken seriously or not.*

Lies differ from true statements in their emotional aspect. Specific statements are usually boring and fragile, whereas specific lies are often sensational and violent. In contrast, general statements are often daring and fragile whereas general lies are usually boring and violent. The truth is fragile. True statements require the author to stick his neck out to criticism. Researchers know that if one doesn't stick the neck out, it can't be an observation or a theory. (A *theory* is another name for one or several connected, not yet falsified universal statements about observations.)** Telling the truth does make vulnerable. For this reason, theories are often *daring*, *arrogant* or *provoking*; at the same time they have to be *fragile* and *vulnerable*. For men, theories thus resemble what they think about women. Darwin's *The origin of the species*, which developed daring theories, illustrates the stark contrast between the numerous boring and solid facts that Darwin collected and the daring theory that he deduced. Boredom of facts is a sign of truth.

In contrast, the witch-hunters propagating 'creationism' or so-called 'intelligent design' are examples of liars. The specific lies they propagate, such as 'the world was created in October 4004 BCE', are sensational, whereas the general lies they propagate, such as 'there have not been big changes in the past', are boring. This is in full contrast with common sense. Moreover, lies, in contrast to true statements, make people violent. The worse the lie, the more violent the people. This connection can be observed regularly in the news. In other words, 'creationism' and 'intelligent design' are not only lies, they are bad lies. A 'good' *general lie*, like a good physical theory, seems crazy and seems vulner-

---

* To clarify the vocabulary usage of this text: *religion* is spirituality plus a varying degree of power abuse. The mixture depends on each person's history, background and environment. *Spirituality* is the open participation in the whole of nature. Most, maybe all, people with a passion for physics are spiritual. Most are not religious.

** In other words, a set of not yet falsified patterns of observations on the same topic is called a *(physical) theory*. The term 'theory' will always be used in this sense in this walk, i.e., with the meaning 'set of correct general statements'. This use results from its Greek origin: 'theoria' means 'observation'; its original meaning, 'passionate and emphatic contemplation', summarizes the whole of physics in a single word. ('Theory', like 'theatre', is formed from the root θέ, meaning 'the act of contemplating'.) Sometimes, however, the term 'theory' is used – being confused with 'hypothesis' – with the meaning of 'conjecture', as in 'your theory is wrong', sometimes with the meaning of 'model', as in 'Chern–Simons' theory and sometimes with the meaning of 'standard procedure', as in 'perturbation theory'. These incorrect uses are avoided here. To bring the issue to a point: the *theory of evolution* is not a conjecture, but a set of correct statements based on observation.

able, such as 'people have free will'. A 'good' *specific lie* is boring, such as 'this looks like bread, but for the next ten minutes it is not'. Good lies do not induce violence. Feelings can thus be a criterion to judge the quality of lies, if we pay careful attention to the type of statement. A number of common lies are discussed later in this chapter.

An important aspect of any 'good' lie is to make as few *public* statements as possible, so that critics can check as little as possible. (For anybody sending corrections of mistakes in this text, I provide a small reward.) To detect lies, public scrutiny is important, though not always reliable. Sometimes, even scientists make statements which are not based on observations. However, a 'good' lie is always well prepared and told on purpose; accidental lies are frowned upon by experts. Examples of good lies in science are 'aether', 'UFOs', 'creation science', or 'cold fusion'. Sometimes it took many decades to detect the lies in these domains.

To sum up, the central point of the art of lying without being caught is simple: do not divulge details. Be *vague*. All the methods used to verify a statement ask for details, for *precision*. For any statement, its degree of precision allows one to gauge the degree to which the author is sticking his neck out. The more precision that is demanded, the weaker a statement becomes, and the more likely a fault will be found, if there is one. This is the main reason that we chose an increase in precision as a guide for our mountain ascent. By the way, the same method is used in criminal trials. To discover the truth, investigators typically ask all the witnesses a large number of questions, allowing as many *details* as possible come to light. When sufficient details are collected, and the precision is high enough, the situation becomes clear. Telling 'good' lies is much more difficult than telling the truth; it requires an excellent imagination.

> “ Truth is an abyss. ”
>
> Democritus

> “ To teach superstitions as truth is a most terrible thing. ”
>
> Hypatia of Alexandria (*c.* 355–415)

> “ [Absolute truth:] It is what scientists say it is when they come to the end of their labors. ”
>
> Charles Peirce

IS THIS STATEMENT TRUE? – A BIT ABOUT NONSENSE

> “ There are three types of people: those who believe in Father Christmas, those who do not believe in Father Christmas, and those who are Father Christmas. ”
>
> Anonymous

> “ Truth is a rhetorical concept. ”
>
> Paul Feyerabend

Not all statements can be categorized as true or false. There is a third option: statements can simply make no sense. There are even such statements in mathematics, where they are called *undecidable*. An example is the continuum hypothesis. This hypothesis is undecidable because it makes a statement that depends on the precise meaning of the term

'set'; in standard mathematical usage the term is not defined sufficiently precisely so that a truth value can be assigned to the continuum hypothesis. In short, statements can be undecidable because the concepts contained in them are not sharply defined.

Statements can also be undecidable for other reasons. Phrases such as 'This statement is not true' illustrate the situation. Kurt Gödel* has even devised a general way of constructing such statements in the domain of logic and mathematics. The different variations of these *self-referential* statements, especially popular both in the field of logic and computer science, have captured a large public.** Similarly undecidable statements can be constructed with terms such as 'calculable', 'provable' and 'deducible'.

In fact, self-referential statements are undecidable because they are meaningless. If the usual definition of 'true', namely corresponding to facts, is substituted into the sentence 'This statement is not true', we quickly see that it has no meaningful content. The most famous meaningless sentence of them all was constructed by the linguist Noam Chomsky:

Colorless green ideas sleep furiously.

It is often used as an example for the language processing properties of the brain, but nobody sensible elevates it to the status of a paradox and writes philosophical discussions about it. To do that with the title of this section is a similar waste of energy.

The main reason for the popular success of self-reference is the difficulty in perceiving the lack of meaning.*** A good example is the statement:

This statement is false or you are an angel.

We can actually deduce from it that 'you are an angel.' Can you see how? If you want, you can change the second half and get even more interesting statements. Such examples show that statements referring to themselves have to be treated with great care when under investigation. In short, whenever you meet somebody who tries to use the self-referential construction by Kurt Gödel to deduce another statement, take a step back, or better, a few more. Self-reference, especially the type defined by Gödel, is a hard but common path – especially amongst wannabe-intellectuals – to think, tell and write *nonsense*. Nothing useful can be deduced from nonsense. Well, not entirely; it does help to meet psychiatrists on a regular basis.

---

* Kurt Gödel (1906–1978), famous Austrian logician.
** A general introduction is given in the beautiful books by RAYMOND SMULLYAN: *Satan, Cantor and Infinity and Other Mind-boggling Puzzles*, Knopf, 1992; *What is the Name of This Book? The Riddle of Dracula and Other Logical Puzzles*, Touchstone, 1986, and *The Lady or the Tiger? And Other Puzzles*, Times Books, 1982. Also definitions can have no content, such as David Hilbert's 'smallest number that has not been mentioned this century' or 'the smallest sequence of numbers that is described by more signs than this sentence'.
*** A well-known victim of this difficulty is Paulus of Tarsus. The paradox of the Cretan poet *Epimenedes* (6th century BCE) who said 'All Cretans lie' is too difficult for the notoriously humour-impaired Paulus, who in his letter to Titus (chapter 1, verses 12 and 13, in the christian bible) calls Epimenedes a 'prophet', adds some racist comments, and states that this 'testimony' is true. But wait! There is a final twist to this story. The statement 'All Cretans lie' is *not* a paradox at all; a truth value can actually be ascribed to it, because the statement is not really self-referential. Can you confirm this? The only *genuine* paradox is 'I am lying', to which it is indeed impossible to ascribe a truth value.

In physics, in the other natural sciences, and in legal trials these problems do not emerge, because self-referential statements are not used.* In fact, the work of logicians confirms, often rather spectacularly, that there is no way to extend the term 'truth' beyond the definition of 'correspondence with facts.'

> Ein Satz kann unmöglich von sich selbst
> aussagen, daß er wahr ist.**
> Ludwig Wittgenstein, *Tractatus*, 4.442

CURIOSITIES AND FUN CHALLENGES ABOUT LIES AND NONSENSE

> A man is his own easiest dupe, for what he
> wishes to be true he generally believes to be
> true.
> Demosthenes, 349 BCE.

Ref. 213

> Quator vero sunt maxima comprehendendae
> veritatis offendicula, quae omnem
> quemcumque sapientem impediunt, et vix
> aliquem permittunt ad verum titulum
> sapientiae pervenire: videlicet fragilis et
> indignae auctoritatis exemplum, consuetudinis
> diurnitatis, vulgi sensus imperiti, et propriae
> ignorantiae occultatio cum ostentatione
> sapientiae apparentis.***
> Roger Bacon, *Opus majus*, 1267.

> Es ist ja nicht alles, was ich den Bürgern sage,
> gelogen.****
> Konrad Adenauer, 1962, West German
> Chancellor.

Some lies are entertaining and funny – and are better called jokes –, some are signs of psychic disturbance, and some are made with criminal intent. Some statements are not lies, but simply nonsense. Have fun distinguishing them.

* *

Challenge 252 e 'Yesterday I drowned.' Is this a good or a bad lie?

* *

During a church sermon, a man fell asleep. He dreamt about the French revolution: he was being brought to the guillotine. At that moment, his wife noticed that he was sleeping. In the same moment in which the man dreamt that the knife was hitting him, his wife gave him a tap on his neck with her fan. The shock instantly killed the man. – Is this story true or false?

Challenge 253 e

Challenge 251 ny
* Why are circular statements, like those of Galilean physics, not self-referential?
** 'It is quite impossible for a proposition to state that it itself is true.'
*** 'There are four stumbling blocks to truth and knowledge: weak and unworthy authority, custom, popular prejudice, and the concealment of ignorance with apparent knowledge.'
**** 'Indeed, not everything that I tell the people is a lie.'

* *

Starting in the 1990s, so-called *crop circles* were produced by people walking with stilts, a piece of wood and some rope into fields of crops. Nevertheless, many pretended and even more believed that these circles were made by extraterrestrial beings. Is this a good or a bad lie? Can you find some reasons why this is impossible?

Challenge 254 ny

* *

Often one hears or reads statements like: 'mind (or spirit or soul) is stronger than matter.' Beware of anybody who says this; he wants something from you. Can you show that these statements are all and always wrong?

Challenge 255 e

* *

In certain countries, two lies were particularly frequent in the early twenty-first century. The first: global warming does not exist. The second: global warming is not due to human causes. Are these good or bad lies?

Challenge 256 s

* *

Sometimes it is heard that a person whose skin is completely covered with finest metal powder will die, due to the impossibility of the skin to breathe. Can you show from you own observation that this is wrong?

Challenge 257 s

* *

A famous mixture of hoax and belief premises that the Earth was created about six thousand years ago. (Some believers even use this false statement as justification for violence against non-believers.) Can you explain why the number is wrong?

Challenge 258 ny

* *

A famous provocation: the world has been created last Saturday. Can you decide whether this is wrong?

Challenge 259 s

* *

Hundreds of hoaxes are found on the www.museumofhoaxes.com website. It gives an excellent introduction into the art of lying; of course it exposes only those who have been caught. Enjoy the science stories, especially those about archaeology. (Several other sites with similar content can be found on the internet.)

* *

In the 1990s, many so-called 'healers' in the Philippines earned large amounts of money by suggesting patients that they were able to extract objects from their bodies without operating. Why is this not possible? (For more information on health lies, see the www.quackwatch.com website.)

Challenge 260 e

* *

Since the 1980s, people have claimed that it is possible to acquire knowledge simply from somebody 1000 km away, without any communication between the two people. However, the assumed 'morphogenetic fields' cannot exist. Why not?

Challenge 261 e

* *

It is claimed that a Fire Brigade building in a city in the US hosts a light bulb that has been burning without interruption since 1901 (at least this was the case in 2005). Can this be true? Hundreds of such stories, often called 'urban legends,' can be found on the www.snopes.com website. However, some of the stories are not urban legends, but true, as the site shows.

Challenge 262 s

* *

'This statement has been translated from French into English.' Is the statement true, false or neither?

* *

Aeroplanes have no seat row 13. Many tall hotels have no floor 13. What is the lie behind this habit? What is the truth behind it?

Challenge 263 s

* *

For about a thousand years, certain people pretend that they have been stigmatized, i.e., that they have 'miraculously' suffered wounds that are similar to those of Jesus's crucifixion. How can one prove by a one-second observation that all of these people, without exception, produced the wounds by themselves?

Challenge 264 s

* *

'In the middle age and in antiquity, people believed in the flat Earth.' This is a famous lie that is rarely questioned. The historian Reinhard Krüger has shown that the lie is most of all due to the writers Thomas Paine (1794) and Washington Irving (1928). Fact is that since Aristotle, everybody believed in a spherical Earth.

* *

Challenge 265 s   Is the term 'multiverse', a claimed opposite to 'universe', a lie or a belief?

* *

The following is not a lie. A good way to suppress curiosity in children is used in many environments: let the child watch television whenever it wants. Do it for a few weeks and you will not recognize the child any more. Do it for a few years, and the curiosity will not come back at all.

* *

Challenge 266 e   How would you show that 'Earth rays' are a lie?

* *

How would you show that the statement 'the laws of nature could change any time' is a lie?

Challenge 267 s

* *

Challenge 268 e   'I can generate energy from the vacuum.' Show that this is a lie.

* *

'Not everything that exists can be measured.' Show that this is a lie.

* *

'Not everything is known.' This statement is quite interesting: modern physics indeed claims the opposite in many domains. For example, all forms of energy are known; so are all forms of moving entities. In short, even though this statement is correct – indeed, not everything is known – it is often used by liars. Be careful when you hear it.

* *

Here is a lie that uses mathematics, from a journalist: 'Your university exams treat women applicants worse than men; your statistics show that only 41 % of all female, but 57% of all male applicants are admitted.' The university is small and has only two faculties; so it checks its numbers.

Faculty 1 admitted 60% of all males (60 of 100 applicants) and 65% of all applicant females (13 of 20 applicants). Faculty 2 admitted 30% of all males (3 of 10 applicants) and 32% of all females (16 of 50 applicants).

In total, the university thus admitted 63 of 110 male applicants (or 57%) and 29 of 70 female applicants (or 41%). In other words, even though in each faculties the percentage of admitted females was *higher*, the *total* admission percentage for females was *lower*.
Why? In fact, this is a true story; in this version, the numbers are simplified, to make the situation as clear as possible. But a large university once got in trouble with journalists in this way, despite preferring women in each of its departments. Some journalists are excellent liars.

* *

A domain in which lies are common is the food industry. It is now possible to buy artificial eggs, artificial tomato, or artificial shrimps. Many bread products contain cysteine; for many decades, cysteine was extracted from human hair!

* *

The British Broadcasting Corporation is famous for its April 1st pranks. One of the best
ever is its documentary on flying penguins. Search for it on the internet.

OBSERVATIONS

> “ Knowledge is a sophisticated statement of
> ignorance.
>   Attributed to Karl Popper ”

The collection of a large number of true statements about a type of observations, i.e., of a large number of facts, is called *knowledge*. Where the domain of observations is
sufficiently extended, one speaks of a *science*. A *scientist* is thus somebody who collects

knowledge.* We found above that an observation is classified input into the memory of several people. Since there is motion all around, to describe all these observations is a mammoth task. As for every large task, to a large extent the use of appropriate tools determines the degree of success that can be achieved. These tools, in physics and in all other sciences, fall in three groups: tools for the collection of observations, tools to communicate observations and tools to communicate relations between observations. The latter group has been already discussed in the section on language and on mathematics. We just touch on the other two.

HAVE ENOUGH OBSERVATIONS BEEN RECORDED?

> Every generation is inclined to define 'the end of physics' as coincident with the end of their scientific contributions.
>
> Julian Schwinger**

Physics is an experimental science; it rests on the collection of observations. To realize this task effectively, all sorts of *instruments*, i.e., tools that facilitate observations, have been developed and built. Microscopes, telescopes, oscilloscopes, as well as thermometers, hygrometers, manometers, pyrometers, spectrometers amongst others are familiar examples. The precision of many of these tools is being continuously improved even today; their production is a sizeable part of modern industrial activity, examples being electrical measuring apparatus and diagnostic tools for medicine, chemistry and biology. Instruments can be as small as a tip of a few tungsten atoms to produce an electron beam of a few volts, and as large as 27 km in circumference, producing an electron beam with more than 100 GV effective accelerating voltage. Instruments have been built that contain and measure the coldest known matter in the universe. Other instruments can measure length variations of much less than a proton diameter over kilometre long distances. Instruments have been put deep inside the Earth, on the Moon, on several planets, and have been sent outside the solar system.

In this walk, instruments are not described; many good textbooks on this topic are available. Most observations collected by instruments are not mentioned here. The most important results in physics are recorded in standard publications, such as the Landolt–Börnstein series and the physics journals (Appendix E gives a general overview of information sources).

Ref. 216, Ref. 217

Ref. 218

---

* The term 'scientist' is a misnomer peculiar to the English language. Properly speaking, a 'scientist' is a follower of *scientism*, an extremist philosophical school that tried to resolve all problems through science. For this reason, some religious sects have the term in their name. Since the English language did not have a shorter term to designate 'scientific persons', as they used to be called, the term 'scientist' started to appear in the United States, from the eighteenth century onwards. Nowadays the term is used in all English-speaking countries – but not outside them, fortunately.

** Julian Seymour Schwinger (1918–1994), US-American infant prodigy. He was famous for his clear thinking and his excellent lectures. He worked on waveguides and synchroton radiation, made contributions to nuclear physics and developed quantum electrodynamics. For the latter he received the 1965 Nobel Prize in Physics together with Tomonaga and Feynman. He was a thesis advisor to many famous physicists and wrote several excellent and influential textbooks. Nevertheless, at the end of his life, he became strangely interested in a hoax turned sour: cold fusion.

Ref. 215

Will there be significant new future observations in the domain of the fundamentals of motion? At present, *in this specific domain*, even though the number of physicists and publications is at an all-time high, the number of new experimental discoveries has been steadily diminishing for many years and is now fairly small. The sophistication and investment necessary to obtain new results has become extremely high. In many cases, measuring instruments have reached the limits of technology, of budgets or even those of nature. The number of new experiments that produce results showing no deviation from theoretical predictions is increasing steadily. The number of historical papers that try to enliven dull or stalled fields of enquiry are increasing. Claims of new effects which turn out to be false, due to measurement errors, self-deceit or even fraud have become so frequent that scepticism has become a common response. Although in many domains of science, including physics, discoveries are still expected, on the fundamentals of motion the arguments just presented seem to show that new observations are only a remote possibility. The task of collecting observations on the foundations of motion (though not on other topics of physics) seems to be *complete*. Indeed, most observations described here were obtained before the end of the twentieth century. We are not too early with our walk.

> "Measure what is measurable; make measurable what is not.
>
> Wrongly attributed to Galileo.

Are all physical observables known?

> "Scientists have odious manners, except when you prop up their theory; then you can borrow money from them.
>
> Mark Twain

The most practical way to communicate observations was developed a long time ago: by measurements. A measurement allows effective communication of an observation to other times and places. This is not always as trivial as it sounds; for example, in the Middle Ages people were unable to compare precisely the 'coldness' of the winters of two different years! The invention of the thermometer provided a reliable solution to this requirement. A *measurement* is thus the classification of an observation into a standard set of observations; to put it simply, a measurement is a *comparison with a standard*. This definition of a measurement is precise and practical, and has therefore been universally adopted. For example, when the length of a house is measured, this aspect of the house is classified into a certain set of standard lengths, namely the set of lengths defined by multiples of a unit. A *unit* is the abstract name of the standard for a certain observable. Numbers and units allow the most precise and most effective communication of measurement results.

For all measurable quantities, practical standard units and measurement methods have been defined; the main ones are listed and defined in Appendix A. All units are derived from a few fundamental ones; this is ultimately due to our limited number of senses: length, time and mass are related to sight, hearing and touch. Our limited number of senses is, in turn, due to the small number of observables of nature.

We call *observables* the different measurable aspects of a system. Most observables,

such as size, speed, position, etc. can be described by numbers, and in this case they are *quantities*, i.e., multiples of some standard unit. Observables are usually abbreviated by *(mathematical) symbols*, usually letters from some alphabet. For example, the symbol *c* commonly specifies the velocity of light. For most observables, standard symbols have been defined by international bodies.* The symbols for the observables that describe the state of an object are also called *variables*. Variables on which other observables depend are often called *parameters*. (Remember: a parameter is a variable constant.) For example, the speed of light is a constant, the position a variable, the temperature is often a parameter, on which the length of an object, for example, can depend. Note that not all observables are quantities; in particular, parities are not multiples of any unit.

Today the task of defining tools for the communication of observations can be considered *complete*. (For quantities, this is surely correct; for parity-type observables there could be a few examples to be discovered.) This is a simple and strong statement. Even the BIPM, the Bureau International des Poids et Mesures, has stopped adding new units.**

As a note, the greatness of a physicist can be ranked by the number of observables he has introduced. Even a great scientist such as Einstein, who discovered many 'laws' of nature, only introduced one new observable, namely the metric tensor for the description of gravity. Following this criterion – as well as several others – Maxwell is the most important physicist, having introduced electric and magnetic fields, the vector potential, and several other material dependent observables. For Heisenberg, Dirac and Schrödinger, the wave function describing electron motion could be counted as half an observable (as it is a quantity necessary to calculate measurement results, but not itself an observable). Incidentally, even the introduction of *any* term that is taken up by others is a rare event; 'gas', 'entropy' and only a few others are such examples. It has always been much more difficult to discover an observable than to discover a 'law'; usually, observables are developed by many people cooperating together. Indeed, many 'laws' bear people's names, but almost no observables.

If the list of observables necessary to describe nature is complete, does this mean that all the patterns or rules of nature are known? No; in the history of physics, observables were usually defined and measured long *before* the precise rules connecting them were found. For example, all observables used in the description of motion itself, such as time, position and its derivatives, momentum, energy and all the thermodynamic quantities, were defined before or during the nineteenth century, whereas the most precise versions of the patterns or 'laws' of nature connecting them, special relativity and non-equilibrium thermodynamics, have been found only in the twentieth century. The same is true for all observables connected to electromagnetic interaction. The corresponding patterns of nature, quantum electrodynamics, was discovered long after the corresponding observables. The observables that were discovered last were the fields of the strong and the weak nu-

---

* All mathematical symbols used in this walk, together with the alphabets from which they are taken, are listed in Appendix A on notation. They follow international standards whenever they are defined. The standard symbols of the physical quantities, as defined by the International Standards Organization (ISO), the International Union of Pure and Applied Physics (IUPAP) and the International Union of Pure and Applied Chemistry (IUPAC), can be found for example in the *bible*, i.e., the *CRC Handbook of Chemistry and Physics*, CRC Press, Boca Raton, 1992.

** The last, the katal or mol/s, was introduced in 1999. Physical units are presented in Appendix A.

clear interactions. Also, in this case, the patterns of nature were formulated much later.*

DO OBSERVATIONS TAKE TIME?

An observation is an interaction with some part of nature leading to the production of a record, such as a memory in the brain, data on a tape, ink on paper, or any other fixed pattern applied to a support. The necessary irreversible interaction process is often called *writing* the record. Obviously, writing takes a certain amount of time; zero interaction time would give no record at all. Therefore any recording device, including our brain, always records some *time average* of the observation, however short it may be.

What we call a fixed image, be it a mental image or a photograph, is always the time average of a moving situation. Without time averaging, we would have no fixed memories. On the other hand, any time averaging introduces a blur that hides certain details; and in our quest for precision, at a certain moment, these details are bound to become important. The discovery of these details will begin in the upcoming part of the walk, the one centred on quantum theory. In the final part of our mountain ascent we will discover that there is a shortest possible averaging time. Observations of that short duration show so many details that even the distinction between particles and empty space is lost. In contrast, our concepts of everyday life appear only after relatively long time averages. The search for an average-free description of nature is one of the big challenges of our adventure.

IS INDUCTION A PROBLEM IN PHYSICS?

> " Nur *gesetzmäßige* Zusammenhänge sind *denkbar*.**
>
> Ludwig Wittgenstein, *Tractatus*, 6.361

> " There is a tradition of opposition between adherents of induction and of deduction. In my view it would be just as sensible for the two ends of a worm to quarrel.
>
> Alfred North Whitehead

Induction is the usual term used for the act of making, from a small and finite number of experiments, general conclusions about the outcome of *all* possible experiments performed in other places, or at other times. In a sense, it is the technical term for sticking out one's neck, which is necessary in every scientific statement. Induction has been a major topic of discussion for science commentators. Frequently one finds the remark that knowledge in general, and physics in particular, relies on induction for its statements. According to some, induction is a type of hidden belief that underlies all sciences but at the same time contrasts with them.

To avoid wasting energy, we make only a few remarks. The first can be deduced from a simple experiment. Try to convince a critic of induction to put their hand into a fire.

---

* Is it possible to talk about observations at all? It is many a philosopher's hobby to discuss whether there actually is an example for an 'Elementarsatz' – an atomic fact – mentioned by Wittgenstein in his *Tractatus*. There seems to be at least one that fits: *Differences exist*. It is a simple sentence; in the final part of our walk, it will play a central role.

** 'Only connexions that are *subject to law* are *thinkable*.'

Nobody who honestly calls induction a belief should conclude from a few unfortunate experiences in the past that such an act would also be dangerous in the future... In short, somehow induction works.

A second point is that physical universal statements are always openly stated; they are never hidden. The refusal to put one's hand into a fire is a consequence of the invariance of observations under time and space translations. Indeed, general statements of this type form the very basis of physics. However, no physical statement is a belief only because it is universal; it always remains open to experimental checks. Physical induction is not a hidden method of argumentation, it is an explicit part of experimental statements. In fact, the complete list of 'inductive' statements used in physics is well known.. These statements are so important that they have been given a special name: they are called *symmetries*. The list of all known symmetries of nature is the candidate list for all inductive statements used in physics.

Perhaps the best argument for the use of induction is that there is no way to avoid it when one is thinking. There is no way to think, to talk or to remember without using concepts, i.e., without assuming that most objects or entities have the same properties over time. There is also no way to communicate with others without assuming that the observations made from the other's viewpoint are similar to one's own. There is no way to think without symmetry and induction. Indeed, the concepts related to symmetry and induction, such as space and time, belong to the fundamental concepts of language. The only sentences which do not use induction, the sentences of logic, do not have any content (*Tractatus*, 6.11). Indeed, without induction, we cannot classify observations at all! Evolution has given us memory and a brain because induction works. To criticize induction is not to criticize natural sciences, it is to criticize the use of thought in general. We should never take too seriously people who themselves do what they criticize in others; sporadically pointing out the ridicule of this endeavour is just the right amount of attention they deserve.

The topic could be concluded here, were it not for some interesting developments in modern physics that put two additional nails in the coffin of arguments against induction. First, in physics whenever we make statements about all experiments, all times or all velocities, such statements are actually about a *finite number* of cases. We know today that infinities, both in size and in number, do not occur in nature. The infinite number of cases appearing in statements in classical physics and in quantum mechanics are apparent, not real, and due to human simplifications and approximations. Statements that a certain experiment gives the same result 'everywhere' or that a given equation is correct for 'all times', always encompass only a *finite* number of examples. A great deal of otherwise often instinctive repulsion to such statements is avoided in this way. In the sciences, as well as in this book, 'all' *never* means an infinite number of cases.

Finally, it is well known that extrapolating from a few cases to many is false when the few cases are independent of each other. However, this conclusion is correct if the cases are interdependent. From the fact that somebody found a penny on the street on two subsequent months, cannot follow that he will find one the coming month. Induction is only correct if we know that all cases have similar behaviour, e.g. because they follow from the same origin. For example, if a neighbour with a hole in his pocket carries his salary across that street once a month, and the hole always opens at that point because of the beginning of stairs, then the conclusion would be correct. It turns out that the results of modern physics encountered in the final part of our walk show that all situations in nature are indeed interdependent, and thus we prove in detail that what is called 'induction' is in fact a logically correct conclusion.

> " In the progress of physics, the exception often
> turned out to be the general case. "

THE QUEST FOR PRECISION AND ITS IMPLICATIONS

> " Der Zweck der Philosophie ist die logische
> Klärung der Gedanken.*
> Ludwig Wittgenstein, *Tractatus*, 4.112 "

To talk well about motion means to talk precisely. Precision requires avoiding three common mistakes in the description of nature.

First, concepts should never have a contradiction built into their definition. For example, any phenomenon occurring in nature evidently is a 'natural' phenomenon; therefore, to talk about either 'supernatural' phenomena or 'unnatural' phenomena is a mistake that nobody interested in motion should let go unchallenged; such terms contain a logical contradiction. Naturally, *all* observations are natural. Incidentally, there is a reward of more than a million dollars for anybody proving the opposite. In over twenty years, nobody has yet been able to collect it.

Ref. 220

Second, concepts should not have unclear or constantly changing definitions. Their content and their limits must be kept constant and explicit. The opposite of this is often encountered in crackpots or populist politicians; it distinguishes them from more reliable thinkers. Physicists can also fall into the trap; for example, there is, of course, only one *single* (physical) universe, as even the name says. To talk about more than one universe is an increasingly frequent error.

Ref. 221

Third, concepts should not be used outside their domain of application. It is easy to succumb to the temptation to transfer results from physics to philosophy without checking the content. An example is the question: 'Why do particles follow the laws of nature?' The flaw in the question is due to a misunderstanding of the term 'laws of nature' and to a confusion with the laws of the state. If nature were governed by 'laws', they could be changed by parliament. Remembering that 'laws of nature' simply means 'pattern', 'property' or 'description of behaviour', and rephrasing the question correctly as 'Why do particles behave in the way we describe their behaviour?' one can recognize its senselessness.

---

* 'The object of philosophy is the logical clarification of thoughts.'

In the course of our walk, we will often be tempted by these three mistakes. A few such situations follow, with the ways of avoiding them.

> Consistency is the last refuge of the unimaginative.
>
> Oscar Wilde

What are interactions? – No emergence

> The whole is always more than the sum of its parts.
>
> Aristotle, *Metaphysica*, 10f–1045a.

In the physical description of nature, the whole is always *more* than the sum of its parts. Actually, the difference between the whole and the sum of its parts is so important that it has a special name: the *interaction* between the parts. For example, the energy of the whole minus the sum of the energies of its parts is called the energy of interaction. In fact, the study of interactions is the main topic of physics. In other words, physics is concerned *primarily* with the difference between the parts and the whole, contrary to what is often suggested by bad journalists or other sloppy thinkers.

Note that the term 'interaction' is based on the general observation that anything that affects anything else is, in turn, affected by it; interactions are *reciprocal*. For example, if one body changes the momentum of another, then the second changes the momentum of the first by the same (negative) amount. The reciprocity of interactions is a result of conservation 'laws'. The reciprocity is also the reason that somebody who uses the term 'interaction' is considered a heretic by monotheistic religions, as theologians regularly point out. They repeatedly stress that such a reciprocity implicitly denies the immutability of the deity. (Are they correct?)

The application of the definition of interaction also settles the frequently heard question of whether in nature there are 'emergent' properties, i.e., properties of systems that cannot be deduced from the properties of their parts and interactions. By definition, there are no emergent properties. 'Emergent' properties can only appear if interactions are approximated or neglected. The idea of 'emergent' properties is a product of minds with restricted horizons, unable to see or admit the richness of consequences that general principles can produce. In defending the idea of emergence, one belittles the importance of interactions, working, in a seemingly innocuous, maybe unconscious, but in fact sneaky way, against the use of reason in the study of nature. 'Emergence' is a belief.

The simple definition of interaction given above sounds elementary, but it leads to surprising conclusions. Take the atomic idea of Democritus in its modern form: nature is made of vacuum and of particles. The first consequence is the *paradox of incomplete description*: experiments show that there are interactions between vacuum and particles. However, interactions are differences between parts and the whole, in this case between vacuum and particles on the one hand, and the whole on the other. We thus have deduced that nature is not made of vacuum and particles alone.

The second consequence is the *paradox of overcomplete description*: experiments also show that interactions happen through exchange of particles. However, we have counted particles already as basic building blocks. Does this mean that the description of nature by vacuum and particles is an overdescription, counting things twice?

We will resolve both paradoxes in the last part of our mountain ascent.

WHAT IS EXISTENCE?

> " You know what I like most? Rhetorical
> questions. "

Ref. 223   Assume a friend tells you 'I have seen a *grampus* today!' You would naturally ask what it looks like. What answer do we expect? We expect something like 'It's an animal with a certain number of heads similar to a *X*, attached to a body like a *Y*, with wings like a *Z*, it make noises like a *U* and it felt like a *V*' – the letters denoting some other animal or object. Generally speaking, in the case of an object, this scene from Darwin's voyage to South America shows that in order to talk to each other, we first need certain basic, common concepts ('animal', 'head', 'wing', etc.). In addition, for the definition of a new entity we need a characterization of its parts ('size', 'colour'), of the way these parts relate to each other, and of the way that the whole interacts with the outside world ('feel', 'sound'). In other words, for an object to exist, we must be able to give a list of relations with the outside world. An object exists if we can interact with it. (Is observation sufficient to

Challenge 275 s   determine existence?)

For an abstract concept, such as 'time' or 'superstring', the definition of existence has to be refined only marginally: *(physical) existence is the effectiveness to describe interactions accurately*. This definition applies to trees, time, virtual particles, imaginary numbers, entropy and so on. It is thus pointless to discuss whether a physical concept 'exists' or whether it is 'only' an abstraction used as a tool for descriptions of observations. The two possibilities coincide. The point of dispute can only be whether the description provided by a concept is or is not *precise*.

For mathematical concepts, existence has a somewhat different meaning: a mathematical concept is said to exist if it has no built-in contradictions. This is a much weaker requirement than physical existence. It is thus incorrect to deduce physical existence from mathematical existence. This is a frequent error; from Pythagoras' times onwards it was often stated that since mathematical concepts exist, they must therefore also exist in nature. Historically, this error occurred in the statements that planet orbits 'must' be circles, that planet shapes 'must' be spheres or that physical space 'must' be Euclidean. Today this is still happening with the statements that space and time 'must' be continuous and that nature 'must' be described by sets. In all these cases, the reasoning is wrong. In fact, the continuous attempts to deduce physical existence from mathematical existence hide that the opposite is correct: a short reflection shows that mathematical existence is a special

Challenge 276 s   case of physical existence.

We note that there is also a different type of existence, namely *psychological existence*. A concept can be said to exist psychologically if it describes human internal experience. Thus a concept can exist psychologically even if it does not exist physically. It is easy to

Challenge 277 s   find examples from the religions or from systems that describe inner experiences. Also myths, legends and comic strips define concepts that only exist psychologically, not physically. In our walk, whenever we talk about existence, we mean physical existence only.

Do things exist?

> "Wer Wissenschaft und Kunst besitzt,
> Hat auch Religion;
> Wer jene beiden nicht besitzt,
> Der habe Religion.*
> Johann Wolfgang von Goethe, *Zahme Xenien,*
> *IX*

Using the above definition of existence, the question becomes either trivial or imprecise. It is trivial in the sense that things necessarily exist if they describe observations, since they were defined that way. But perhaps the questioner meant to ask: Does reality exist independently of the observer?

Using the above, this question can be rephrased: 'Do the things we observe exist independently of observation?' After thousands of years of extensive discussion by professional philosophers, logicians, sophists and amateurs the answer is the same: it is 'Yes', because the world did not change after great-grandmother died. The disappearance of observers does not seem to change the universe. These experimental findings can be corroborated by inserting the definition of 'existence' into the question, which then becomes: 'Do the things we observe interact with other aspects of nature when they do not interact with people?' The answer is evident. Several popular books on quantum mechanics fantasize about the importance of the 'mind' of observers – whatever this term may mean; they provide pretty examples of authors who see themselves as irreplaceable, seemingly having lost the ability to see themselves as part of a larger entity.

Of course there are other opinions about the existence of things. The most famous is that of the Irishman George Berkeley (1685–1753) who rightly understood that thoughts based on observation alone, if spread, would undermine the basis of the religious organization of which he was one of the top managers. To counteract this tendency, in 1710 he published *A Treatise Concerning the Principles of Human Knowledge*, a book denying the existence of the material world. This reactionary book became widely known in likeminded circles (it was a time when few books were written) even though it is based on a fundamentally flawed idea: it assumes that the concept of 'existence' and that of 'world' can be defined independently. (You may be curious to try the feat.)

Berkeley had two aims when he wrote his book. First, he tried to deny the capacity of people to arrive at judgements on nature or on any other matter *from their own experience*. Second, he also tried to deny the *ontological reach* of science, i.e., the conclusions one can draw from experience on the questions about human existence. Even though Berkeley is generally despised nowadays, he actually achieved his main aim: he was the originator of the statement that science and religion do not contradict, but *complement* each other. By religion, Berkeley did not mean either morality or spirituality; every scientist is a friend of both of these. By religion, Berkeley meant that the standard set of beliefs for which he stood is above the deductions of reason. This widely cited statement, itself a belief, is still held dearly by many even to this day. However, when searching for the origin of motion, all beliefs stand in the way, including this one. Carrying beliefs is like carrying oversized baggage: it prevents one from reaching the top of Motion Mountain.

Challenge 278 e

---

* He who possesses science and art, also has religion; he who does not possess the two, better have religion.

DOES THE VOID EXIST?

> *Natura abhorret vacuum.*
>
> Antiquity

> Teacher: 'What is found between the nucleus
> and the electrons?'
> Student: 'Nothing, only air.'

In philosophical discussions '*void*' is usually defined as 'non-existence'. It then becomes a game of words to ask for a yes or no answer to the question 'Does the void exist?' The expression 'the existence of non-existence' is either a contradiction of terms or is at least unclearly defined; the topic would not seem to be of great interest. However, similar questions do appear in physics, and a physicist should be prepared to notice the difference of this from the previous one. Does a vacuum exist? Does empty space exist? Or is the world 'full' everywhere, as the more conservative biologist Aristotle maintained? In the past, people have been killed for giving an answer that was unacceptable to authorities.

It is not obvious, but it is nevertheless important, that the modern physical concepts of 'vacuum' and 'empty space' are not the same as the philosophical concept of 'void'. 'Vacuum' is not defined as 'non-existence'; on the contrary, it is defined as the absence of matter and radiation. Vacuum is an entity with specific observable properties, such as its number of dimensions, its electromagnetic constants, its curvature, its vanishing mass, its interaction with matter through curvature and through its influence on decay, etc. (A table of the properties of a physical vacuum is given on page 112.) Historically, it took a long time to clarify the distinction between a physical vacuum and a philosophical void. People confused the two concepts and debated the existence of the vacuum for more than two thousand years. The first to state that it existed, with the courage to try to look through the logical contradiction at the underlying physical reality, were Leucippus and Democritus, the most daring thinkers of antiquity. Their speculations in turn elicited the reactionary response of Aristotle, who rejected the concept of vacuum. Aristotle and his disciples propagated the belief about nature's *horror of the vacuum*.

The discussion changed completely in the seventeenth century, when the first experimental method to realize a vacuum was devised by Torricelli.* Using mercury in a glass tube, he produced the first laboratory vacuum. Can you guess how? Arguments against the existence of the vacuum again appeared around 1900, when it was argued that light needed 'aether' for its propagation, using almost the same arguments that had been used two hundred years earlier, but in different words. However, experiments failed to detect any of the supposed properties of this unclearly defined concept. Experiments in the field of general relativity showed that a vacuum can move – though in a completely different way from the way in which the aether was expected to move – that the vacuum can be bent, but it then tends to return to its shape. Then, in the late twentieth century, quantum field theory again argued against the existence of a true vacuum and in favour of a space full of virtual particle–antiparticle pairs, culminating in the discussions around the cosmological constant.

---

* Evangelista Torricelli (b. 1608 Faenza, d. 1647 Florence), Italian physicist, pupil and successor to Galileo. The (non-SI) pressure unit 'torr' is named after him.

The question 'Does the void exist?' is settled conclusively only in the last part of this walk, in a rather surprising way.

Is nature infinite?

> It is certain and evident to our senses, that in the world some things are in motion. Now whatever is moved is moved by another... If that by which it is moved be itself moved, then this also needs to be to be moved by another, and that by another again. But this cannot go on to infinity, because then there would be no first mover and consequently, no other mover, seeing that subsequent movers move only inasmuch as they are moved by the first mover, as the staff moves only because it is moved by the hand. Therefore it is necessary to arrive at a first mover, moved by no other; and this everyone understands to be god.
>
> Thomas Aquinas (*c.* 1225–1274) *Summa Theologiae*, I, q. 2.

Most of the modern discussions about set theory centre on ways to defining the term 'set' for various types of infinite collections. For the description of motion this leads to two questions: Is the universe infinite? Is it a set? We begin with the first one. Illuminating the question from various viewpoints, we will quickly discover that it is both simple and imprecise.

Do we need infinite quantities to describe nature? Certainly, in classical and quantum physics we do, e.g. in the case of space-time. Is this necessary? We can say already a few things.

Any set can be finite in one aspect and infinite in another. For example, it is possible to proceed along a finite mathematical distance in an infinite amount of time. It is also possible to travel along any distance whatsoever in a given amount of mathematical time, making infinite speed an option, even if relativity is taken into account, as was explained earlier.

Despite the use of infinities, scientists are still limited. We saw above that many types of infinities exist. However, no infinity larger than the cardinality of the real numbers plays a role in physics. No space of functions or phase space in classical physics and no Hilbert space in quantum theory has higher cardinality. Despite the ability of mathematicians to define much larger kinds of infinities, the description of nature does not need them. Even the most elaborate descriptions of motion use only the infinity of the real numbers.

But is it possible at all to say *of nature* or of one of its aspects that it is indeed *infinite*? Can such a statement be compatible with observations? No. It is evident that every statement that claims that something in nature is infinite is a belief, and is not backed by observations. We shall patiently eliminate this belief in the following.

The possibility of introducing false infinities make any discussion on whether humanity is near the 'end of science' rather difficult. The amount of knowledge and the time required to discover it are unrelated. Depending on the speed with which one advances

through it, the end of science can be near or unreachable. In practice, scientists have thus the power to *make* science infinite or not, e.g. by reducing the speed of progress. As scientists need funding for their work, one can guess the stand that they usually take.

In short, the universe cannot be proven to be infinite. But can it be *finite*? At first sight, this would be the only possibility left. (It is not, as we shall see.) But even though many have tried to describe the universe as finite in all its aspects, no one has yet been successful. In order to understand the problems that they encountered, we continue with the other question mentioned above:

Is the universe a set?

A simple observation leads us to question whether the universe is a set. For 2500 years
it has been said that the universe is made of vacuum and particles. This implies that the universe is made of a certain *number* of particles. Perhaps the only person to have taken this conclusion to the limit was the English astrophysicist Arthur Eddington (1882–1944), who wrote:

> 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.

Eddington was ridiculed over and over again for this statement and for his beliefs that lead up to it. His arguments were indeed based on his personal preferences for certain pet numbers. However, we should not laugh too loudly. In fact, for 2500 years almost all scientists have thought along the same line, the only difference being that they have left the precise number unspecified! In fact, *any other number* put into the above sentence would be equally ridiculous. Avoiding specifying it is just a coward's way of avoiding looking at this foggy aspect of the particle description of nature.

Is there a particle number at all in nature? If you smiled at Eddington's statement, or if you shook your head over it, it may mean that you instinctively believe that nature is not a set. Is this so? Whenever we define the universe as the totality of events, or as the totality of all space-time points and objects, we imply that space-time points can be distinguished, that objects can be distinguished and that both can be distinguished from each other. We thus assume that nature is separable and a set. But is this correct? The question is important. The ability to distinguish space-time points and particles from each other is often called *locality*. Thus the universe is separable or is a set if and only if our description of it is local.* And in everyday life, locality is observed without exception.

In daily life we also observe that nature is separable and a whole at the same time. It is a 'many that can be thought as one': in daily life nature is a set. Indeed, the basic characteristic of nature is its diversity. In the world around us we observe changes and differences; we observe that nature is separable. Furthermore, all aspects of nature belong together: there are relations between these aspects, often called 'laws,' stating that the different aspects of nature form a whole, usually called the universe.

---

* In quantum mechanics also other, more detailed definitions of locality are used. We will mention them in the quantum part of this text. The issue mentioned here is a different, more fundamental one, and not connected with that of quantum theory.

In other words, the possibility of describing observations with the help of 'laws' follows from our experience of the separability of nature. The more precisely the separability is specified, the more precisely the 'laws' can be formulated. Indeed, if nature were not separable or were not a unity, we could not explain why stones fall downwards. Thus we are led to speculate that we should be able to deduce all 'laws' from the fact that nature is separable.

In addition, only the separability allows us to describe nature at all. A description is a classification, that is, a mapping between certain aspects of nature and certain concepts. All concepts are sets and relations. Since the universe is separable, it can be described with the help of sets and relations. Both are separable entities with distinguishable parts. A precise description is commonly called an understanding. In short, the universe is comprehensible only because it is separable.

Moreover, only the separability of the universe makes our brain such a good instrument. The brain is built from a large number of connected components, and only the brain's separability allows it to function. In other words, thinking is only possible because nature is separable.

Finally, only the separability of the universe allows us to distinguish reference frames, and thus to define all symmetries at the basis of physical descriptions. And in the same way that separability is thus necessary for *covariant* descriptions, the unity of nature is necessary for *invariant* descriptions. In other words, the so-called 'laws' of nature are based on the experience that nature is both separable and unifiable – that it is a set.

These arguments seem overwhelmingly to prove that the universe is a set. However, these arguments apply only to everyday experience, everyday dimensions and everyday energies. Is nature a set also *outside* the domains of daily life? Are objects different at all energies, even when they are looked at with the highest precision possible? We have three open issues left: the issue of the number of particles in the universe; the circular definition of space, time and matter; and the issue as to whether describing nature as made of particles and void is an overdescription, an underdescription, or neither. These three issues make us doubt whether objects are countable at all energies. We will discover in the final part of our mountain ascent that this is not the case in nature. The consequences will be extensive and fascinating. As an example, try to answer the following: if the universe is not a set, what does that mean for space and time?

Does the universe exist?

> " Each progressive spirit is opposed by a
> thousand men appointed to guard the past. "
> Maurice Maeterlink

Following the definition above, existence of a concept means its usefulness to describe interactions. There are two common definitions of the concept of 'universe'. The first is the totality of all matter, energy and space-time. But this usage results in a strange consequence: since nothing can interact with this totality, we cannot claim that the universe exists.

So let us take the more restricted view, namely that the universe is only the totality of all matter and energy. But also in this case it is impossible to interact with the universe. Can you give a few arguments to support this?

In short, we arrive at the conclusion that the universe does not exist. We will indeed confirm this result in more detail later on in our walk. In particular, since the universe does not exist, it does not make sense to even try to answer *why* it exists. The best answer might be: because of furiously sleeping, colourless green ideas.

What is creation?

> (Gigni) De nihilo nihilum, in nihilum nil posse reverti.*
>
> Persius, *Satira*, III, v. 83-84.

> Anaxagoras, discovering the ancient theory that nothing comes from nothing, decided to abolish the concept of creation and introduced in its place that of discrimination; he did not hesitate to state, in effect, that all things are mixed to the others and that discrimination produces their growth.
>
> *Anonymous fragment*, Middle Ages.

The term 'creation' is often heard when talking about nature. It is used in various contexts with different meanings.

One speaks of creation as the characterization of human actions, such as observed in an artist painting or a secretary typing. Obviously, this is one type of change. In the classification of change introduced at the beginning of our walk, the changes cited are movements of objects, such as the electrons in the brain, the molecules in the muscles, the material of the paint, or the electrons inside the computer. This type of creation is thus a special case of motion.

One also speaks of creation in the biological or social sense, such as in 'the creation of life', or 'creation of a business', or 'the creation of civilization'. These events are forms of growth or of self-organization; again, they are special cases of motion.

Physicists one often say that a lamp 'creates' light or that a stone falling into a pond 'creates' water ripples. Similarly, they talk of 'pair creation' of matter and antimatter. It was one of the important discoveries of physics that all these processes are special types of motion, namely excitation of fields.

In popular writing on cosmology, 'creation' is also a term commonly applied, or better misapplied, to the *big bang*. However, the expansion of the universe is a pure example of motion, and contrary to a frequent misunderstanding, the description of the big bang contains only processes that fall into one of the previous three categories, as shown in the relevant chapter in general relativity. The big bang is not an example of creation. Quantum cosmology provides more reasons to support the fact that the naive term 'creation' is not applicable to the big bang. First, it turns out that the big bang was not an event. Second, it was not a beginning. Third, it did not provide a *choice* from a large set of possibilities. The big bang does not have any properties attributed to the term 'creation'.

In summary, we conclude that in all cases, *creation is a type of motion*. (The same applies to the notions of 'disappearance' and 'annihilation'.) No other type of creation is observed in nature. In particular, the naive sense of 'creation', namely 'appearance from

---

* Nothing (can appear) from nothing, nothing can disappear into nothing.

nothing' – *ex nihilo* in Latin – is never observed in nature. All observed types of 'creation' require space, time, forces, energy and matter for their realization. Creation requires something to exist already, in order to take place. In addition, precise exploration shows that no physical process and no example of motion has a beginning. Our walk will show us that nature does not allow us to pinpoint beginnings. This property alone is sufficient to show that 'creation' is not a concept applicable to what happens in nature. Worse still, creation is applied only to physical systems; we will discover that nature is not a system and that systems do not exist.

The opposite of creation is *conservation*. The central statements of physics are conservation theorems: for energy, mass, linear momentum, angular momentum, charge, etc. In fact, every conservation 'law' is a detailed and accurate rejection of the concept of creation. The ancient Greek idea of atoms already contains this rejection. Atomists stated that there is no creation and no disappearance, but only motion of atoms. Every transformation of matter is a motion of atoms. In other words, the idea of the atom was a direct consequence of the negation of creation. It took humanity over 2000 years before

it stopped locking people in jail for talking about atoms, as had happened to Galileo.

However, there is one exception in which the naive concept of creation does apply: it describes what magicians do on stage. When a magician makes a rabbit appear from nowhere, we indeed experience 'creation' from nothing. At its best such magic is a form of entertainment, at its worst, a misuse of gullibility. The idea that the universe results from either of these two does not seem appealing; on second thought though, maybe looking at the universe as the ultimate entertainment could open up a fresh and more productive approach to life.

Voltaire (1694–1778) popularized an argument against creation often used in the past: we do not know whether creation has taken place or not. Today the situation is different: we *do* know that it has *not* taken place, because creation is a type of motion and, as we will see in the concluding part of our mountain ascent, motion did not exist near the big bang.

Have you ever heard the expression 'creation of the laws of nature'? It is one of the most common examples of disinformation. First of all, this expression confuses the 'laws' with nature itself. A description is not the same as the thing itself; everybody knows that giving their beloved a description of a rose is different from giving an actual rose. Second, the expression implies that nature is the way it is because it is somehow 'forced' to follow the 'laws' – a rather childish and, what is more, incorrect view. And third, the expression assumes that it is possible to 'create' descriptions of nature. But a 'law' is a description, and a description by definition cannot be created: so the expression makes no sense at all. The expression 'creation of the laws of nature' is the epitome of confused thinking.

It may well be that calling a great artist 'creative' or 'divine', as was common during the Renaissance, is not blasphemy, but simply an encouragement to the gods to try to do as well. In fact, whenever one uses the term 'creation' to mean anything other than some form of motion, one is discarding both observations and human reason. It is one of the last pseudo-concepts of our modern time; no expert on motion should forget this. It is impossible to escalate Motion Mountain without getting rid of 'creation'. This is not

easy. We will encounter the next attempt to bring back creation in the study of quantum theory.

> “ Every act of creation is first of all an act of
> destruction.   ”
>   Pablo Picasso

IS NATURE DESIGNED?

> “ In the beginning the universe was created. This
> has made a lot of people very angry and has
> been widely regarded as a bad move.   ”
>   Douglas Adams, *The Restaurant at the End of*
>   *the Universe.*

The tendency to infer the creation of an object from its simple existence is widespread. Some people jump to this conclusion every time they see a beautiful landscape. This habit stems from the triple prejudice that a beautiful scene implies a complex description, in turn implying complex building instructions, and therefore pointing to an underlying *design*.

Ref. 228

Vol. I, page 318

This chain of thought contains several mistakes. First, in general, beauty is not a consequence of complexity. Usually it is the opposite: indeed, the study of chaos and of self-organization demonstrates how beautifully complex shapes and patterns can be generated with extremely simple descriptions. True, for most human artefacts, complex descriptions indeed imply complex building processes; a personal computer is a good example of a complex object with a complex production process. But in nature, this connection does not apply. We have seen above that even the amount of information needed to construct a human body is about a million times smaller than the information stored in the brain alone. Similar results have been found for plant architecture and for many other examples of patterns in nature. The simple descriptions behind the apparent complexities of nature have been and are still being uncovered by the study of self-organization, chaos, turbulence and fractal shapes. In nature, complex structures derive from *simple* processes. Beware of anyone who says that nature has 'infinite' or 'high complexity': first of all, complexity is not a measurable entity, despite many attempts to quantify it. In addition, all known complex system can be described by (relatively) few parameters and simple equations. Finally, nothing in nature is infinite.

The second mistake in the argument for design is to link a description with an 'instruction', and maybe even to imagine that some unknown 'intelligence' is somehow pulling the strings of the world's stage. The study of nature has consistently shown that there is no hidden intelligence and no instruction behind the processes of nature. An instruction is a list of orders to an executioner. But there are no orders in nature, and no executioners. There are no 'laws' of nature, only descriptions of processes. Nobody is building a tree; the tree is an outcome of the motion of molecules making it up. The genes in the tree do contain information; but no molecule is given any instructions. What seem to be instructions to us are just natural movements of molecules and energy, described by the same patterns taking place in non-living systems. The whole idea of instruction – like that of 'law' of nature – is an ideology, born from an analogy with monarchy or even tyranny, and a typical anthropomorphism.

The third mistake in the argument for design is the suggestion that a complex description for a system implies an underlying design. This is not correct. A complex description only implies that the system has a long evolution behind it. The correct deduction is:

something of large complexity exists; therefore it has *grown*, i.e., it has been transformed through input of (moderate) energy over time. This deduction applies to flowers, mountains, stars, life, people, watches, books, personal computers and works of art; in fact it applies to all objects in the universe. The complexity of our environment thus points out the considerable age of our environment and reminds us of the shortness of our own life.

The lack of basic complexity and the lack of instructions in nature confirm a simple result: there is not a single observation in nature that implies or requires design or creation. On the other hand, the variety and intensity of nature's phenomena fills us with deep awe. The wild beauty of nature shows us how *small* a part of nature we actually are, both in space and in time.* We shall explore this experience in detail. We shall find that remaining open to nature's phenomena in all their overwhelming intensity is central to the rest of our adventure.

> There is a separation between state and church,
> but not yet between state and science.
> Paul Feyerabend

What is a description?

> In theory, there is no difference between theory
> and practice. In practice, there is.

Following standard vocabulary usage, a description of an observation is a list of the details. The above example of the grampus showed this clearly. In other words, a *description* of an observation is the act of categorizing it, i.e., of comparing, by identifying or distinguishing, the observation with all the other observations already made. A description is a classification. In short, *to describe means to see as an element of a larger set.*

A description can be compared to the 'you are here' sign on a city tourist map. Out of a set of possible positions, the 'you are here' sign gives the actual one. Similarly, a description highlights the given situation in comparison with all other possibilities. For example, the formula $a = GM/r^2$ is a description of the observations relating motion to gravity, because it classifies the observed accelerations a according to distance to the central body r and to its mass M; indeed such a description sees each specific case as an example of a general pattern. The habit of generalizing is one reason for the often disturbing dismissiveness of scientists: when they observe something, their professional training usually makes them classify it as a special case of a known phenomenon and thus keeps them from being surprised or from being exited about it.

A description is thus the opposite of a *metaphor*; the latter is an analogy relating an observation with another *special* case; a description relates an observation with a *general* case, such as a physical theory.

---

* The search for a 'sense' in life or in nature is a complicated (and necessary) way to try to face the smallness of human existence.

> “ Felix qui potuit rerum cognoscere causas,
>   atque metus omnis et inexorabile fatum
>   subjecit pedibus strepitumque acherontis avari. ”
>   Vergilius*

REASON, PURPOSE AND EXPLANATION

> “ Der ganzen modernen Weltanschauung liegt
>   die Täuschung zugrunde, daß die sogenannten
>   Naturgesetze die Erklärungen der
>   Naturerscheinungen seien.** ”
>   Ludwig Wittgenstein, *Tractatus*, 6.371

Compare the following two types of questions and answers:

1. Why are the leaves of most trees green? Because they absorb red and blue light. Why do they absorb those colours? Because they contain chlorophyll. Why is chlorophyll green? Because all chlorophyll types contain magnesium between four pyrrole groups, and this chemical combination gives the green colour, as a result of its quantum mechanical energy levels. Why do plants contain chlorophyll? Because this is what land plants can synthesize. Why only this? Because all land plants originally evolved from the green algae, who are only able to synthesize this compound, and not the compounds found in the blue or in the red algae, which are also found in the sea.

2. Why do children climb trees, and why do some people climb mountains? Because of the sensations they experience during their activity: the feelings of achievement, the symbolic act to go upwards, the wish to get a wider view of the world are part of this type of adventure.

The two types of 'why'-questions show the general difference between reasons and purposes (although the details of these two terms are not defined in the same way by everybody). A *purpose* or *intention* is a classification applied to the actions of humans or animals; strictly speaking, it specifies the quest for a feeling, namely for achieving some type of satisfaction after completion of the action. On the other hand, a *reason* is a specific relation of a fact with the rest of the universe, usually its past. What we call a reason always rests outside the observation itself, whereas a purpose is always internal to it.

Reasons and purposes are the two possibilities of explanations, i.e., the two possible answers to questions starting with 'why'. Usually, physics is not concerned with purpose or with people's feeling, mainly because its original aim, to talk about motion with precision, does not seem to be achievable in this domain. Therefore, *physical* explanations of facts are never purposes, but are always reasons. A *physical explanation* of an observation is always the description of its relation with the rest of nature.***

Ref. 229

---

* 'Happy he who can know the causes of things and who, free of all fears, can lay the inexorable fate and the noise of Acheron to his feet.' *Georgica*, book II, verses 490 ss.) Publius Vergilius Maro (70–19 BCE), the great roman poet, is author of the Aeneid. Acheron was the river crossed by those who had just died and were on their way to the Hades.
** 'The whole modern conception of the world is founded on the illusion that the so-called laws of nature are the explanations of natural phenomena.'
*** It is important to note that purposes are *not* put aside because they pertain to the future, but because they are inadmissible anthropomorphisms. In fact, for deterministic systems, we can equally say that the future is actually a *reason* for the present and the past, a fact often forgotten.

This means that – contrary to common opinion – a question starting with 'why' is accessible to physical investigation as long as it asks for a reason and not for a purpose. In particular, questions such as 'why do stones fall downwards and not upwards?' or 'why do electrons have that value of mass, and why do they have mass at all?' or 'why does space have three dimensions and not thirty-six?' can be answered, as these ask for the connection between specific observations and more general ones. Of course, not all demands for explanation have been answered yet, and there are still problems to be solved. Our present trail only leads from a few answers to some of the more fundamental questions about motion.

The most general quest for an explanation derives from the question: why is the universe the way it is? The topic is covered in our mountain ascent using the two usual approaches, namely:

Unification and demarcation

> "Tout sujet est un; et, quelque vaste qu'il soit, il
> peut être renfermé dans un seul discours.*
> Buffon, *Discours sur le style*. "

Studying the properties of motion, constantly paying attention to increase the accuracy of description, we find that explanations are generally of two types:**

1. 'It is like all such cases; also this one is described by ...' The situation is recognized as a *special case* of a general behaviour.
2. 'If the situation were different, we would have a conclusion in contrast with observations.' The situation is recognized as the *only possible case*.***

In other words, the first approach is to formulate rules or 'laws' that describe larger and larger numbers of observations, and compare the observation with them. This endeavour is called the *unification* of physics – by those who like it; those who don't like it, call it 'reductionism'. For example, the same rule describes the flight of a tennis ball, the motion of the tides at the sea shore, the timing of ice ages, and the time at which the planet Venus ceases to be the evening star and starts to be the morning star. These processes are all consequences of universal gravitation. Similarly, it is not evident that the same rule describes the origin of the colour of the eyes, the formation of lightning, the digestion of food and the working of the brain. These processes are described by quantum electrodynamics.

Unification has its most impressive successes when it predicts an observation that has not been made before. A famous example is the existence of antimatter, predicted

---

* Every subject is one and, however vast it is, it can be comprised in a single discourse.

** Are these the only possible ones?

*** These two cases have not to be confused with similar sentences that *seem* to be explanations, but that aren't:

— 'It is like the case of ...' A similarity with another *single* case is *not* an explanation.
— 'If it were different, it would contradict the idea that ...' A contradiction with an *idea* or with a theory is *not* an explanation.

by Dirac when he investigated the solutions of an equation that describes the precise behaviour of common matter.

The second procedure in the search for explanations is the elimination of all other imaginable alternatives in favour of the actually correct one. This endeavour has no commonly accepted name: it could be called the *demarcation* of the 'laws' of physics – by those who like it; others call it 'anthropocentrism', or simply 'arrogance'.

When we discover that light travels in such a way that it takes the shortest possible time to its destination, when we describe motion by a principle of least action, or when we discover that trees are branched in such a way that they achieve the largest effect with the smallest effort, we are using a demarcation viewpoint.

In summary, unification, answering 'why' questions, and demarcation, answering 'why not' questions, are typical for the progress throughout the history of physics. We can say that the dual aspects of unification and demarcation form the composing and the opposing traits of physics. They stand for the desire to *know everything*.

However, neither demarcation nor unification can explain the universe. Can you see why? In fact, apart from unification and demarcation, there is a third possibility that merges the two and allows one to say more about the universe. Can you find it? Our walk will automatically lead to it later.

PIGS, APES AND THE ANTHROPIC PRINCIPLE

> Das wichtigste Hilfsmittel des Wissenschaftlers ist der Papierkorb.*
>
> Several authors

The wish to achieve demarcation of the patterns of nature is most interesting when we follow the consequences of different rules of nature until we find them in contradiction with the most striking observation: our own human existence. In this special case the program of demarcation is often called the *anthropic principle* – from the Greek ἄνθρωπος, meaning 'man'.

For example, if the Sun–Earth distance were different from what it is, the resulting temperature change on the Earth would have made impossible the appearance of life, which needs liquid water. Similarly, our brain would not work if the Moon did not circle the Earth. It is also well-known that if there were fewer large planets in the solar system, the evolution of humans would have been impossible. The large planets divert large numbers of comets, preventing them from hitting the Earth. The spectacular collision of comet Shoemaker–Levy-9 with Jupiter, the astronomical event of July 1994, was an example of this diversion of a comet.**

Also the anthropic principle has its most impressive successes when it predicts unknown observations. The most famous example stems from the study of stars. Carbon atoms, like all other atoms except most hydrogen, helium or lithium atoms, are formed in stars through fusion. While studying the mechanisms of fusion in 1953, the well-known British astrophysicist Fred Hoyle*** found that carbon nuclei could not be formed from

---

* 'The most important instrument of a scientist is the waste paper basket.'
** For a collection of pictures of this event, see e.g. the garbo.uwasa.fi/pc/gifslevy.html website.
*** Fred Hoyle (b. 1915 Bingley, d. 2001), important British astronomer and astrophysicist. He was the first

the alpha particles present inside stars at reasonable temperatures, unless they had an excited state with an increased cross-section. From the fact of our existence, which is based on carbon, Hoyle thus predicted the existence of a previously unknown excited state of the carbon nucleus. And, indeed, the excited state was found a few months later by Willy Fowler.*

In its *serious* form, the anthropic principle is therefore the quest to deduce the description of nature from the experimental fact of our own existence. In the popular literature, however, the anthropic principle is often changed from a simple experimental method to deduce the patterns of nature, to its *perverted* form, a melting pot of absurd metaphysical ideas in which everybody mixes up their favourite beliefs. Most frequently, the experimental observation of our own existence has been perverted to reintroduce the idea of 'design', i.e., that the universe has been constructed with the aim of producing humans; often it is even suggested that the anthropic principle is an *explanation* – a gross example of disinformation.

How can we distinguish between the serious and the perverted form? We start with an observation. We would get exactly the same rules and patterns of nature if we used the existence of pigs or monkeys as a starting point. In other words, if we would reach *different* conclusions by using the *porcine principle* or the *simian principle*, we are using the perverted form of the anthropic principle, otherwise we are using the serious form. (The carbon-12 story is thus an example of the serious form.) This test is effective because there is no known pattern or 'law' of nature that is particular to humans but unnecessary for apes or pigs.**

> Er wunderte sich, daß den Katzen genau an den
> Stellen Löcher in den Pelz geschnitten wären,
> wo sie Augen hätten.
>
> Georg Christoph Lichtenberg***

Does one need cause and effect in explanations?

> There are in nature neither rewards nor
> punishments – there are only consequences.
>
> Robert Ingersoll

> The world owes you nothing. It was there first.
>
> Mark Twain

---

and maybe only physicist who ever made a specific prediction – namely the existence of an excited state of the carbon nucleus – from the simple fact that humans exist. A permanent maverick, he coined the term 'big bang' even though he did not accept the evidence for it, and proposed another model, the 'steady state'. His most important and well-known research was on the formation of atoms inside stars. He also propagated the belief that life was brought to Earth from extraterrestrial microbes.

* William A. Fowler (1911–1995) shared the 1983 Nobel Prize in Physics with Subramanyan Chandrasekhar for this and related discoveries.

** Though apes do not seem to be good physicists, as described in the text by D. J. Povinelli, *Folk Physics for Apes: the Chimpanzee's Theory of How the World Works*, Oxford University Press, 2000.

*** 'He was amazed that cats had holes cut into their fur precisely in those places where they had eyes.' Georg Christoph Lichtenberg (1742–1799), German physicist and intellectual, professor in Göttingen, still famous today for his extremely numerous and witty aphorisms and satires. Among others of his time, Lichtenberg made fun of all those who maintained that the universe was made exactly to the measure of man, a frequently encountered idea in the foggy world of the anthropic principle.

> No matter how cruel and nasty and evil you
> may be, every time you take a breath you make
> a flower happy.
>
> Mort Sahl

Ref. 232   Historically, the two terms 'cause' and 'effect' have played an important role in philosophical discussions. In particular, during the birth of modern mechanics, it was important to point out that every effect has a cause, in order to distinguish precise thought from thought based on beliefs, such as 'miracles', 'divine surprises' or 'evolution from nothing'. It was equally essential to stress that effects are different from causes; this distinction avoids pseudo-explanations such as the famous example by Molière where the doctor explains to his patient in elaborate terms that sleeping pills work because they contain a dormitive virtue.

But in physics, the concepts of cause and effect are not used at all. That miracles do not appear is expressed every time we use symmetries and conservation theorems. The observation that cause and effect differ from each other is inherent in any evolution equation. Moreover, the concepts of cause and effect are not clearly defined; for example, it is especially difficult to define what is meant by one cause as opposed to several of them, and the same for one or several effects. Both terms are impossible to quantify and to measure. In other words, useful as 'cause' and 'effect' may be in personal life for distinction between events that regularly succeed each other, they are not necessary in physics. In physical explanations, they play no special roles.

> Ἀγαθὸν καὶ ξαξόν · ἔν καὶ ταὐτό.*
>
> Heraclitus

> Wenn ein Arzt hinter dem Sarg seines Patienten
> geht, so folgt manchmal tatsächlich die Ursache
> der Wirkung.**
>
> Robert Koch

Is consciousness required?

> Variatio delectat.***
>
> Cicero

Ref. 233   A lot of mediocre discussions are going on about this topic, and we will skip them here. What is consciousness? Most simply and concretely, consciousness means the possession of a small part of oneself that is watching what the rest of oneself is perceiving, feeling, thinking and doing. In short, consciousness is the ability to observe oneself, and in particular one's inner mechanisms and motivations. *Consciousness* is the ability of introspection. For this reason, consciousness is *not* a prerequisite for studying motion. Indeed, animals, plants or machines are also able to observe motion. For the same reason, consciousness is not necessary to observe quantum mechanical motion. On the other hand,

---

* 'Good and bad – one and the same.'
** 'When a doctor walks behind the coffin of his patient, indeed the cause sometimes follows the effect.'
*** 'Change pleases.' Marcus Tullius Cicero (106–43 bce), important lawyer, orator and politician at the end of the Roman republic.

both the study of motion and that of oneself have a lot in common: the need to observe carefully, to overcome preconceptions, to overcome fear and the fun of doing so.

For the time being, we have put enough emphasis on the precision of concepts. Talking about motion is also something to be deeply enjoyed. Let us see why.

> " Precision and clarity obey the indeterminacy relation: their product is constant. "
>
> Niels Bohr

Curiosity

> " Precision is the child of curiosity. "

Like the history of every person, also the history of mankind charts a long struggle to avoid the pitfalls of accepting the statements of authorities as truth, without checking the facts. Indeed, whenever curiosity leads us to formulate a question, there are always two general ways to proceed. One is to check the facts personally, the other is to ask somebody. However, the last way is dangerous: it means to give up a part of oneself. Healthy people, children whose curiosity is still alive, as well as scientists, choose the first way. After all, science is adult curiosity.

Curiosity, also called the *exploratory drive*, plays strange games with people. Starting with the original experience of the world as a big 'soup' of interacting parts, curiosity can drive one to find *all* the parts and *all* the interactions. It drives not only people. It has been observed that when rats show curious behaviour, certain brain cells in the hypothalamus get active and secrete hormones that produce positive feelings and emotions. If a rat has the possibility, via some implanted electrodes, to excite these same cells by pressing a switch, it does so voluntarily: rats get *addicted* to the feelings connected with curiosity. Like rats, humans are curious because they enjoy it. They do so in at least four ways: because they are artists, because they are fond of pleasure, because they are adventurers and because they are dreamers. Let us see how.

Originally, curiosity stems from the desire to interact in a positive way with the environment. Young children provide good examples: curiosity is a natural ingredient of their life, in the same way that it is for other mammals and a few bird species; incidentally, the same taxonomic distribution is found for play behaviour. In short, all animals that play are curious, and vice versa. Curiosity provides the basis for learning, for creativity and thus for every human activity that leaves a legacy, such as art or science. The artist and art theoretician Joseph Beuys (1920–1986) had as his own guiding principle that *every* creative act is a form of art. Humans, and especially children, enjoy curiosity because they feel its importance for creativity, and for growth in general.

Curiosity regularly leads one to exclaim: 'Oh!', an experience that leads to the second reason to be curious: relishing feelings of wonder and surprise. Epicurus (Epikuros) (341–271 BCE) maintained that this experience, θαυμάζειν, is the origin of philosophy. These feelings, which nowadays are variously called religious, spiritual, numinous, etc., are the same as those to which rats can become addicted. Among these feelings, Rudolf Otto has introduced the now classical distinction into the fascinating and the frightening. He named the corresponding experiences 'mysterium fascinans' and 'mysterium tremen-

Ref. 234

Ref. 235

dum'.* Within these distinctions, physicists, scientists, children and connoisseurs take a clear stand: they choose the fascinans as the starting point for their actions and for their approach to the world. Such feelings of fascination induce some children who look at the night sky to dream about becoming astronomers, some who look through a microscope to become biologists or physicists, and so on. (It could also be that genetics plays a role in this pleasure of novelty seeking.)

Ref. 236

Perhaps the most beautiful moments in the study of physics are those appearing after new observations have shaken our previously held thinking habits, have forced us to give up a previously held conviction, and have engendered the feeling of being lost. When, in this moment of crisis, we finally discover a more adequate and precise description of the observations, which provide a better insight into the world, we are struck by a feeling usually called illumination. Anyone who has kept alive the memory and the taste for these magic moments knows that in these situations one is pervaded by a feeling of union between oneself and the world.** The pleasure of these moments, the adventures of the change of thought structures connected with them, and the joy of insight following them provides the drive for many scientists. Little talk and lots of pleasure is their common denominator. In this spirit, the great Austrian physicist Victor Weisskopf (1908 –2002) liked to say jokingly: 'There are two things that make life worth living: Mozart and quantum mechanics.'

The choice of moving away from the tremendum towards the fascinans stems from an innate desire, most obvious in children, to reduce uncertainty and fear. This drive is the father of all adventures. It has a well-known parallel in ancient Greece, where the first men studying observations, such as Epicurus, stated explicitly that their aim was to free people from unnecessary fear by deepening knowledge and transforming people from frightened passive victims into fascinated, active and responsible beings. Those ancient thinkers started to popularize the idea that, like the common events in our life, the rarer events also follow rules. For example, Epicurus underlined that lightning is a natural phenomenon caused by interactions between clouds, and stressed that it was a natural process, i.e., a process that followed rules, in the same way as the falling of a stone or any other familiar process of everyday life.

Investigating the phenomena around them, philosophers and later scientists succeeded in freeing people from most of their fears caused by uncertainty and a lack of knowledge about nature. This liberation played an important role in the history of human culture and still pervades in the personal history of many scientists. The aim to arrive at stable, rock-bottom truths has inspired (but also hindered) many of them; Albert Einstein is a well-known example for this, discovering relativity, helping to start up but then denying quantum mechanics.

---

* This distinction is the basis of RUDOLF OTTO, *Das Heilige – Über das Irrationale in der Idee des Göttlichen und sein Verhältnis zum Rationalen*, Beck 1991. This is a new edition of the epoch-making work originally published at the beginning of the twentieth century. Rudolf Otto (1869–1937) was one of the most important theologians of his time.

** Several researchers have studied the situations leading to these magic moments in more detail, notably the Prussian physician and physicist Hermann von Helmholtz (1821–1894) and the French mathematician Henri Poincaré (1854–1912). They distinguish four stages in the conception of an idea at the basis of such a magic moment: saturation, incubation, illumination and verification.

Ref. 237

Interestingly, in the experience and in the development of every human being, curiosity, and therefore the sciences, appears *before* magic and superstition. Magic needs deceit to be effective, and superstition needs indoctrination; curiosity doesn't need either. Conflicts of curiosity with superstitions, ideologies, authorities or the rest of society are thus preprogrammed.

Curiosity is the exploration of limits. For every limit, there are two possibilities: the limit can turn out to be real or apparent. If the limit is real, the most productive attitude is that of acceptance. Approaching the limit then gives strength. If the limit is only apparent and in fact non-existent, the most productive attitude is to re-evaluate the mistaken view, extract the positive role it performed, and then cross the limit. Distinguishing between real and apparent limits is only possible when the limit is investigated with great care, openness and unintentionality. Most of all, exploring limits need courage.

> Das gelüftete Geheimnis rächt sich.*
>
> Bert Hellinger

COURAGE

> Il est dangereux d'avoir raison dans des choses où des hommes accrédités ont tort.**
>
> Voltaire

> Manche suchen Sicherheit, wo Mut gefragt ist, und suchen Freiheit, wo das Richtige keine Wahl läßt.***
>
> Bert Hellinger

Most of the material in this chapter is necessary in the adventure to get to the top of Motion Mountain. But we need more. Like any enterprise, curiosity also requires courage, and complete curiosity, as aimed for in our quest, requires complete courage. In fact, it is easy to get discouraged on this trip. The journey is often dismissed by others as useless, uninteresting, childish, confusing, damaging or, most often, evil. For example, between the death of Socrates in 399 BCE and Paul Thierry, Baron d'Holbach, in the eighteenth century, no book was published with the statement 'gods do not exist', because of the threats to the life of anyone who dared to make the point. Even today, this type of attitude still abounds, as the newspapers show.

Through the constant elimination of uncertainty, both curiosity and scientific activity are implicitly opposed to any idea, person or organization that tries to avoid the comparison of statements with observations. These 'avoiders' demand to live with superstitions and beliefs. But superstitions and beliefs produce unnecessary fear. And fear is the basis of all unjust authorities. One gets into a vicious circle: avoiding comparison with observation produces fear – fear keeps unjust authority in place – unjust authority avoids comparison with observation – etc.

As a consequence, curiosity and science are fundamentally opposed to unjust authority, a connection that made life difficult for people such as Anaxagoras (500–428 BCE)

---

* 'The unveiled secret takes revenge.'
** 'It is dangerous to be right in matters where established men are wrong.'
*** 'Some look for security where courage is required and look for freedom where the right way doesn't
leave any choice.'

in ancient Greece, Hypatia in the Christian Roman empire, Galileo Galilei in the church state, Antoine Lavoisier in France and Albert Einstein (and many others) in Germany. In the second half of the twentieth century, victims were Robert Oppenheimer, Melba Phillips and Chandler Davis in the United States and Andrei Sakharov in the Soviet Union. Each of them tell a horrible but instructive story, as have, more recently, Fang Lizhi, Xu Liangying, Liu Gang and Wang Juntao in China, Kim Song-Man in South Corea, Otanazar Aripov in Uzbekistan, Ramadan al-Hadi al-Hush in Libya, Bo Bo Htun in Burma, Sami Kilani and Salman Salman in Palestine, Abdus Salam in Pakistan, as well as many hundreds of others. In many authoritarian societies the antagonism between curiosity and injustice has hindered or even completely suppressed the development of physics and other sciences, with extremely negative economic, social and cultural consequences.

When embarking on this ascent, we need to be conscious of what we are doing. In fact, external obstacles can be avoided or at least largely reduced by keeping the project to oneself. Other difficulties still remain, this time of personal nature. Many have tried to embark on this adventure with some hidden or explicit intention, usually of an ideological nature, and then have got entangled by it before reaching the end. Some have not been prepared to accept the humility required for such an endeavour. Others were not prepared for the openness required, which can shatter deeply held beliefs. Still others were not ready to turn towards the unclear, the dark and the unknown, confronting them at every occasion.

On the other hand, the dangers are worth it. By taking curiosity as a maxim, facing disinformation and fear with all one's courage, one achieves freedom from all beliefs. In exchange, you come to savour the fullest pleasures and the deepest satisfaction that life has to offer.

We thus continue our hike. At this point, the trail towards the top of Motion Mountain is leading us towards the next adventure: discovering the origin of sizes, shapes and colours in nature.

> And the gods said to man: 'Take what you want, and pay the price.'
>
> Popular saying

> It is difficult to make a man miserable while he feels he is worthy of himself.
>
> Abraham Lincoln

CLASSICAL PHYSICS IN A NUTSHELL

Classical electrodynamics, with mechanics, thermodynamics and relativity, completes our walk through classical physics. In the structure of physics, classical physics encompasses four of the eight points that make up all of physics: mechanics and heat, relativity, gravity and electricity. We have thus covered one half of our adventure. Let summarize what we have found out about motion and what we did not.

What can move?

In nature, *objects*, *radiation* and *space-time* move. In all three cases, motion happens in such a way as to minimize change. Change is also called (physicall) action. In short, all motion minimizes action. And in all three cases of motion, we distinguish the fixed, *intrinsic properties* from the varying *state*. We learned to distinguish and to characterize the possible intrinsic properties and the possible states.

About *objects*: when we searched for all their *fixed, intrinsic* aspects, we found that all sufficiently small objects or particles are described completely by their *mass* and their electric *charge*. There is no magnetic charge. Mass and electric charge are thus the only localized intrinsic properties of classical, everyday objects. Both mass and electric charge are defined by the accelerations they produce around them. Both quantities are conserved; thus they can be added (with certain precautions). Mass, in contrast to charge, is always positive. Mass describes the interaction of objects in collisions and in gravitation, charge the interaction with electromagnetic fields.

All *varying* aspects of objects, i.e., their *state*, can be described using *momentum* and *position*, as well as *angular momentum* and *orientation*. These four quantities can vary continuously in amount and direction. Therefore the set of all possible states forms a space, the so-called *phase space*. The state of extended, shape-changing objects is given by the states of all its constituent particles. These particles make up all objects and somehow interact electromagnetically.

The Lagrangian determines the action, or total change, of any motion. Action, or change, is independent of the observer; the state is not. The states found by different observers are related: the relations are called the 'laws' of motion. For example, for different times they are called *evolution equations*, for different places and orientations they are called *transformation relations*, and for different gauges they are called *gauge transformations*. All the behaviour of the motion of objects is contained in the principle of least action.

Radiation also moves. Everyday types of radiation, such as light, radio waves and their related forms, are travelling electromagnetic waves. They are described by same equations that describe the interaction of charged or magnetic objects. The speed of massless fields is the maximum possible energy speed in nature and is the same for all observers. The motion of radiation describes the motion of images. The *intrinsic properties* of radiation are its dispersion relation and its energy–angular momentum relation. The *state* of radiation is described by its electromagnetic field strength, its phase, its polarization and its coupling to matter. The motion of the electromagnetic field and of radiation minimizes action and change.

Space-time is also able to move, by changing its curvature. The intrinsic properties of space-time are the number of dimensions, its signature and its topology. The state is given by the metric, which describes distances and curvature, and thus the local warpedness. The warpedness can oscillate and propagate, so that empty space can move like a wave. Also the motion of space-time minimizes action and change.

The universe has maximum age and distance values. The universe has a long history, and on large scales, all matter in the universe moves away from all other matter. The large scale topology of space-time is simple, but at this point of our adventure it is unclear what happens at its spatial and temporal limits.

PROPERTIES OF CLASSICAL MOTION

Objects, radiation, and space-time move. Classical physics distilled several specific properties of classical, or everyday motion.

1. Everyday motion is *continuous*. Continuous motion allows us to define space and time. All energy moves in the way space-time dictates it, and space moves the way energy dictates it. This relation describes the motion of the stars, of thrown stones, of light beams and of the tides. Rest and free fall are the same, and gravity is curved space-time. Mass breaks conformal symmetry and thus distinguishes space from time.

   Energy and mass speed is bound from above by a universal constant c, and energy change per time is bound from above by a universal constant $c^5/4G$. The speed value c is realized for the motion of massless particles. It also relates space to time. The power value $c^5/4G$ is realized by horizons. Horizons are found around black holes and at the border of the universe. The maximum power value also relates space-time curvature to energy flow and thus describes the elasticity of space-time.

   No two objects can be at the same spot at the same time. This is the first statement that humans encounter about electromagnetism. It is due to the repulsion of charges of the same sign found in matter. More detailed investigation shows that electric charge accelerates other charges, that charge is necessary to define length and time intervals, and that charges are the source of electromagnetic fields. Also light is such a field. Light travels at the maximum possible velocity c. In contrast to objects, light can interpenetrate.

2. Classical motion *conserves* mass, momentum, energy and angular momentum. Nothing appears out of nothing. Conservation appplies to linear and rotational motion, and to motion of matter, radiation or space-time. Energy and momentum are sim-

ilar to a continuous substances: they are never destroyed, never created, but always redistributed.

3. Classical motion is *relative*: motion depends on the observer.
4. Everyday motion is *reversible*: everyday motion can occur backwards.
5. Everyday motion is *mirror-invariant*: everyday motion can occur in a mirror-reversed way. In short, we found that the classical motion of objects, radiation and space-time is right–left symmetric.
6. Everyday motion is *lazy*: motion happens in a way that minimizes change, i.e., physical action.

In summary, classical physics showed us that motion is *predictable*: nature follows patterns and rules, and there are *no* surprises in nature.

The future of planet Earth

Maybe nature shows no surprises, but it still provides many adventures. On the 8th of March 2002, a 100 m sized body almost hit the Earth. It passed at a distance of only 450 000 km from our planet. On impact, it would have destroyed a region the size of London. Such events occur regularly.* Several other adventures can be predicted by classical physics; they are listed in Table 23. Several items are problems facing humanity in the distant future, but some, such as volcanic eruptions or asteroid impacts, could happen at any time. All are research topics.

Ref. 241

TABLE 23 Examples of disastrous motion of possible future importance.

CRITICAL SITUATION	YEARS FROM NOW
Giant tsunami from volcanic eruption at Canary islands	*c.* 10-200
End of fundamental physics, with a final proof that nature is not infinite	*c.* 20 (around year 2030)
Major nuclear material accident or weapon use	unknown
Explosion of volcano in Greenland, leading to long darkening of sky	unknown
Explosion of Yellowstone or other giant volcano leading to year-long volcanic winter	0 to 100 000
Earth's mantle instability leading to massive volcanic activity	unknown
Mini ice age due to collapse of gulf stream	unknown
Ozone shield reduction	*c.* 100
Rising ocean levels due to greenhouse warming	*c.* 100-1 000
End of applied physics	> 200
Several magnetic north and south poles appear, allowing solar storms to disturb radio and telecommunications, to interrupt electricity supplies, to increase animal mutations and to disorient migrating animals such as wales, birds and tortoises	*c.* 800

* The web pages around cfa-www.harvard.edu/iau/lists/Closest.html provide more information on such events.

TABLE 23 (Continued) Examples of disastrous motion of possible future importance.

CRITICAL SITUATION	YEARS FROM NOW
Our interstellar gas cloud detaches from the solar systems, changing the size of the heliosphere, and thus expose us more to aurorae and solar magnetic fields	$c.$ 3 000
Reversal of Earth's magnetic field, implying a time with almost no magnetic field, with increased cosmic radiation levels and thus more skin cancers and miscarriages	unknown
Atmospheric oxygen depletion due to forest reduction and exaggerated fuel consumption	> 1000
Upcoming ice age	$c.$ 15 000
Possible collision with interstellar gas cloud assumed to be crossed by the Earth every 60 million years, maybe causing mass extinctions	$c.$ 50 000
Possible genetic degeneration of homo sapiens due to Y chromosome reduction	$c.$ 200 000
Africa collides with Europe, transforming the Mediterranean into a lake that starts evaporating	around $3 \cdot 10^6$
Gamma ray burst from within our own galaxy, causing radiation damage to many living beings	between 0 and $5 \cdot 10^6$
Asteroid hitting the Earth, generating tsunamis, storms, darkening sunlight, etc.	between 0 and $50 \cdot 10^6$
Neighbouring star approaching, starting comet shower through destabilization of Oort cloud and thus risk for life on Earth	$> 10^6$
American continent collides with Asia	$> 100 \cdot 10^6$
Molecular cloud engulfs the solar system	unknown
Instability of solar system	$> 100 \cdot 10^6$
Low atmospheric CO_2 content stops photosynthesis	$> 100 \cdot 10^6$
Collision of Milky Way with star cluster or other galaxy	$> 150 \cdot 10^6$
Sun ages and gets hotter, evaporating seas	$> 250 \cdot 10^6$
Ocean level increase due to Earth rotation slowing/stopping (if not evaporated before)	$> 10^9$
Temperature rise/fall (depending on location) due to Earth rotation stop	$> 10^9$
Sun runs out of fuel, becomes red giant, engulfs Earth	$5.0 \cdot 10^9$
Sun stops burning, becomes white dwarf	$5.2 \cdot 10^9$
Earth core solidifies, removing magnetic field and thus Earth's cosmic radiation shield	$10.0 \cdot 10^9$
Nearby nova (e.g. Betelgeuse) bathes Earth in annihilation radiation	unknown
Nearby supernova (e.g. Eta Carinae) blasts over solar system	unknown
Galaxy centre destabilizes rest of galaxy	unknown
Universe recollapses – if ever (see page 122)	$> 20 \cdot 10^9$
Matter decays into radiation – if ever (see Appendix B)	$> 10^{33}$

TABLE 23 (Continued) Examples of disastrous motion of possible future importance.

CRITICAL SITUATION	YEARS FROM NOW
Problems with naked singularities	only in science fiction
Vacuum becomes unstable	only in science fiction

Despite the fascination of the predictions, we leave aside these literally tremendous issues and continue on our adventure.

THE ESSENCE OF CLASSICAL PHYSICS – THE INFINITELY SMALL AND THE LACK OF SURPRISES

We can summarize classical physics with two simple statements: First, *classical physics is the description of motion using the concept of the infinitely small.* Secondly, *nature lacks surprises.*

All concepts used so far, be they for motion, space, time or observables, assume that the infinitely small exists. Special relativity, despite the speed limit, still allows infinitely small velocities; general relativity, despite its black hole limit, still allows infinitely small force and power values. Similarly, in the description of electrodynamics and gravitation, both integrals and derivatives are abbreviations of mathematical processes that use and assume infinitely small distances and time intervals. In other words, the classical description of nature introduces the infinitely small in the description of motion.

Using the infinitely small as a research tool, the classical description of motion discovers that energy, momentum, angular momentum and charge are conserved. They are conserved also for infinitely small dimensions or time intervals. In other words, there are no surprises in motion.

The detailed study of conservation lead us to a strong conclusion: exploring the infinitely small shows that motion is deterministic. The existence of real surprises would contradict determinism.*

Classical physics is the absence of surprises. As reassuring as this result may be, it leaves us with a doubt. Both special and general relativity have eliminated the existence of the infinitely large. There is no infinitely large force, power, size, age or speed. Why should the infinitely small exist, but the infinitely large not? In fact, there are still more open questions about motion.

---

* No surprises also imply no miracles. Classical physics is thus in opposition to many religions. Indeed, many religions argue that infinity is the necessary ingredient to perform miracles. Classical physics shows that this is not the case.

Why have we not yet reached the top of the mountain?

> "The more important fundamental laws and facts
> of physical science have all been discovered,
> and these are now so firmly established that the
> possibility of their ever being supplanted in
> consequence of new discoveries is exceedingly
> remote... Our future discoveries must be looked
> for in the sixth place of decimals."
>
> Albert Michelson, 1894.[*]

We might think that we know nature now, as did Albert Michelson at the end of the nineteenth century. He claimed that electrodynamics and Galilean physics implied that the major laws of physics were well known. The statement is often quoted as an example of flawed predictions, since it reflects an incredible mental closure to the world around him. General relativity was still unknown, and so was quantum theory.

At the end of the nineteenth century, the progress in technology due to the use of electricity, chemistry and vacuum technology had allowed better and better machines and apparatuses to be built. All were built with classical physics in mind. In the years between 1890 and 1920, these classical machines completely destroyed the foundations of classical physics. Experiments with these apparatuses showed that matter is made of atoms, that electrical charge comes in the smallest amounts and that nature behaves randomly. In short, precise experiments show that nature does produce surprises – through in a restricted sense, as we will see. Like the an old empire, the reign of classical physics collapsed. Speaking simply, classical physics does not describe nature correctly at small scales.

But even without machines, many physicists of the time predicted that important changes in the description of nature were necessary. Michelson had overlooked two contradiction between electrodynamics and nature for which he had no excuse. First of all, we found that clocks and metre bars are necessarily made of matter and necessarily based on electromagnetism. But as we just saw, classical electrodynamics does not explain the stability of matter. Matter is made of small particles, but the relation between these particles, electricity and the smallest charges is not clear. If we do not understand matter, we do not yet understand space and time, since space and time are defined using measurement devices made of matter.

Worse, Michelson overlooked a second simple fact: the classical description of nature does not allow us to understand *life*. The abilities of living beings – growing, seeing, hearing, feeling, thinking, being healthy or sick, reproducing and dying – are all unexplained by classical physics. In fact, all these abilities *contradict* classical physics. (Nevertheless, Michelson received the Nobel Prize in Physics in 1907.) Understanding matter and its interactions, including life itself, is therefore the aim of the upcoming part of our ascent of Motion Mountain. In particular, understanding life implies to understand the shape, size, colour and material properties of all things. And this understanding will take place at small scales. To understand nature, we need to study *particles*. Indeed, the atomic structure of matter, the existence of a smallest charge and the existence of a smallest entropy

---

[*] From his address at the dedication ceremony for the Ryerson Physical Laboratory at the University of Chicago.

makes us question the existence of the infinitely small. There is something left to explore. Doing so will lead us from wonder to wonder.
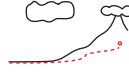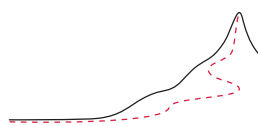
UNITS, MEASUREMENTS AND CONSTANTS

MEASUREMENTS are comparisons with standards. Standards are based on a *unit*. Many different systems of units have been used throughout the world. Most standards confer power to the organization in charge of them. Such power can be misused; this is the case today, for example in the computer industry, and was so in the distant past. The solution is the same in both cases: organize an independent and global standard. For units, this happened in the eighteenth century: to avoid misuse by authoritarian institutions, to eliminate problems with differing, changing and irreproducible standards, and – this is not a joke – to simplify tax collection, a group of scientists, politicians and economists agreed on a set of units. It is called the *Système International d'Unités*, abbreviated *SI*, and is defined by an international treaty, the 'Convention du Mètre'. The units are maintained by an international organization, the 'Conférence Générale des Poids et Mesures', and its daughter organizations, the 'Commission Internationale des Poids et Mesures' and the 'Bureau International des Poids et Mesures' (BIPM), which all originated in the times just before the French revolution.

Ref. 242

SI units

All SI units are built from seven *base units*, whose official definitions, translated from French into English, are given below, together with the dates of their formulation:

- 'The *second* is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium 133 atom.' (1967)*
- 'The *metre* is the length of the path travelled by light in vacuum during a time interval of 1/299 792 458 of a second.' (1983)
- 'The *kilogram* is the unit of mass; it is equal to the mass of the international prototype of the kilogram.' (1901)*
- 'The *ampere* is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed 1 metre apart in vacuum, would produce between these conductors a force equal to $2 \cdot 10^{-7}$ newton per metre of length.' (1948)
- 'The *kelvin*, unit of thermodynamic temperature, is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water.' (1967)*
- 'The *mole* is the amount of substance of a system which contains as many elementary entities as there are atoms in 0.012 kilogram of carbon 12.' (1971)*
- 'The *candela* is the luminous intensity, in a given direction, of a source that emits

monochromatic radiation of frequency $540 \cdot 10^{12}$ hertz and has a radiant intensity in that direction of (1/683) watt per steradian.' (1979)*

Note that both time and length units are defined as certain properties of a standard example of motion, namely light. In other words, also the Conférence Générale des Poids et Mesures makes the point that the observation of motion is a *prerequisite* for the definition and construction of time and space. *Motion is the fundament each observation and measurements.* By the way, the use of light in the definitions had been proposed already in 1827 by Jacques Babinet.*

From these basic units, all other units are defined by multiplication and division. Thus, all SI units have the following properties:

▪ SI units form a system with *state-of-the-art precision*: all units are defined with a precision that is higher than the precision of commonly used measurements. Moreover, the precision of the definitions is regularly being improved. The present relative uncertainty of the definition of the second is around 10^{-14}, for the metre about 10^{-10}, for the kilogram about 10^{-9}, for the ampere 10^{-7}, for the mole less than 10^{-6}, for the kelvin 10^{-6} and for the candela 10^{-3}.

▪ SI units form an *absolute* system: all units are defined in such a way that they can be reproduced in every suitably equipped laboratory, independently, and with high precision. This avoids as much as possible any misuse by the standard-setting organization. (The kilogram, still defined with the help of an artefact, is the last exception to this requirement; extensive research is under way to eliminate this artefact from the definition – an international race that will take a few more years. There are two approaches: counting particles, or fixing $\hbar$. The former can be achieved in crystals, the latter using any formula where $\hbar$ appears, such as the formula for the de Broglie wavelength or that of the Josephson effect.)

▪ SI units form a *practical* system: the base units are quantities of everyday magnitude. Frequently used units have standard names and abbreviations. The complete list includes the seven base units, the supplementary units, the derived units and the admitted units.

The *supplementary* SI units are two: the unit for (plane) angle, defined as the ratio of arc length to radius, is the *radian* (rad). For solid angle, defined as the ratio of the subtended area to the square of the radius, the unit is the *steradian* (sr).

The *derived* units with special names, in their official English spelling, i.e., without capital letters and accents, are:

---

* The respective symbols are s, m, kg, A, K, mol and cd. The international prototype of the kilogram is a platinum–iridium cylinder kept at the BIPM in Sèvres, in France. For more details on the levels of the caesium atom, consult a book on atomic physics. The Celsius scale of temperature θ is defined as: $\theta/°C = T/K - 273.15$; note the small difference with the number appearing in the definition of the kelvin. SI also states: 'When the mole is used, the elementary entities must be specified and may be atoms, molecules, ions, electrons, other particles, or specified groups of such particles.' In the definition of the mole, it is understood that the carbon 12 atoms are unbound, at rest and in their ground state. In the definition of the candela, the frequency of the light corresponds to 555.5 nm, i.e., green colour, around the wavelength to which the eye is most sensitive.

* Jacques Babinet (1794–1874), French physicist who published important work in optics.

NAME	ABBREVIATION	NAME	ABBREVIATION
hertz	$Hz = 1/s$	newton	$N = kg\,m/s^2$
pascal	$Pa = N/m^2 = kg/m\,s^2$	joule	$J = Nm = kg\,m^2/s^2$
watt	$W = kg\,m^2/s^3$	coulomb	$C = As$
volt	$V = kg\,m^2/As^3$	farad	$F = As/V = A^2s^4/kg\,m^2$
ohm	$\Omega = V/A = kg\,m^2/A^2s^3$	siemens	$S = 1/\Omega$
weber	$Wb = Vs = kg\,m^2/As^2$	tesla	$T = Wb/m^2 = kg/As^2 = kg/Cs$
henry	$H = Vs/A = kg\,m^2/A^2s^2$	degree Celsius	°C (see definition of kelvin)
lumen	$lm = cd\,sr$	lux	$lx = lm/m^2 = cd\,sr/m^2$
becquerel	$Bq = 1/s$	gray	$Gy = J/kg = m^2/s^2$
sievert	$Sv = J/kg = m^2/s^2$	katal	$kat = mol/s$

We note that in all definitions of units, the kilogram only appears to the powers of 1, 0 and −1. The final explanation for this fact appeared only recently. Can you try to formulate the reason?

The *admitted* non-SI units are *minute*, *hour*, *day* (for time), *degree* 1° = $\pi/180$ rad, *minute* $1' = \pi/10\,800$ rad, *second* $1'' = \pi/648\,000$ rad (for angles), *litre* and *tonne*. All other units are to be avoided.

All SI units are made more practical by the introduction of standard names and abbreviations for the powers of ten, the so-called *prefixes*:*

POWER	NAME		POWER	NAME		POWER	NAME		POWER	NAME	
10^1	deca	da	10^{-1}	deci	d	10^{18}	Exa	E	10^{-18}	atto	a
10^2	hecto	h	10^{-2}	centi	c	10^{21}	Zetta	Z	10^{-21}	zepto	z
10^3	kilo	k	10^{-3}	milli	m	10^{24}	Yotta	Y	10^{-24}	yocto	y
10^6	Mega	M	10^{-6}	micro	μ	unofficial:					
10^9	Giga	G	10^{-9}	nano	n	10^{27}	Xenta	X	10^{-27}	xenno	x
10^{12}	Tera	T	10^{-12}	pico	p	10^{30}	Wekta	W	10^{-30}	weko	w
10^{15}	Peta	P	10^{-15}	femto	f	10^{33}	Vendekta	V	10^{-33}	vendeko	v
						10^{36}	Udekta	U	10^{-36}	udeko	u

■ SI units form a *complete* system: they cover in a systematic way the complete set of observables of physics. Moreover, they fix the units of measurement for all other sciences

* Some of these names are invented (yocto to sound similar to Latin *octo* 'eight', zepto to sound similar to Latin *septem*, yotta and zetta to resemble them, exa and peta to sound like the Greek words ἑξάκις and πεντάκις for 'six times' and 'five times', the unofficial ones to sound similar to the Greek words for nine, ten, eleven and twelve); some are from Danish/Norwegian (atto from *atten* 'eighteen', femto from *femten* 'fifteen'); some are from Latin (from *mille* 'thousand', from *centum* 'hundred', from *decem* 'ten', from *nanus* 'dwarf'); some are from Italian (from *piccolo* 'small'); some are Greek (micro is from μικρός 'small', deca/deka from δέκα 'ten', hecto from ἑκατόν 'hundred', kilo from χίλιοι 'thousand', mega from μέγας 'large', giga from γίγας 'giant', tera from τέρας 'monster').

Translate: I was caught in such a traffic jam that I needed a microcentury for a picoparsec and that my car's fuel consumption was two tenths of a square millimetre.

as well.

■ SI units form a *universal* system: they can be used in trade, in industry, in commerce, at home, in education and in research. They could even be used by extraterrestrial civilizations, if they existed.

■ SI units form a *coherent* system: the product or quotient of two SI units is also an SI unit. This means that in principle, the same abbreviation, e.g. 'SI', could be used for every unit.

The SI units are not the only possible set that could fulfil all these requirements, but they are the only existing system that does so.*

Since every measurement is a comparison with a standard, any measurement requires *matter* to realize the standard (even for a speed standard), and *radiation* to achieve the comparison. The concept of measurement thus assumes that matter and radiation exist and can be clearly separated from each other.

Challenge 288 e

PRECISION AND ACCURACY OF MEASUREMENTS

Measurements are the basis of physics. Every measurement has an *error*. Errors are due to lack of precision or to lack of accuracy. *Precision* means how well a result is reproduced when the measurement is repeated; *accuracy* is the degree to which a measurement corresponds to the actual value. Lack of precision is due to accidental or *random errors*; they are best measured by the *standard deviation*, usually abbreviated σ; it is defined through

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \ , \tag{102}$$

where $\bar{x}$ is the average of the measurements x_i. (Can you imagine why $n - 1$ is used in the formula instead of n?) For most experiments, the distribution of measurement values tends towards a normal distribution, also called *Gaussian distribution*, whenever the number of measurements is increased. The distribution, shown in Figure 118, is described by the expression

Challenge 289 s

$$N(x) \approx \mathrm{e}^{-\frac{(x-\bar{x})^2}{2\sigma^2}} \ . \tag{103}$$

The square σ^2 of the standard deviation is also called the *variance*. For a Gaussian distribution of measurement values, 2.35σ is the full width at half maximum.

Challenge 290 e

Lack of accuracy is due to *systematic errors*; usually these can only be estimated. This estimate is often added to the random errors to produce a *total* experimental error, sometimes also called *total uncertainty*.

Ref. 245

---

* Apart from international units, there are also *provincial* units. Most provincial units still in use are of Roman origin. The mile comes from *milia passum*, which used to be one thousand (double) strides of about 1480 mm each; today a nautical mile, once defined as minute of arc on the Earth's surface, is exactly 1852 m). The inch comes from *uncia/onzia* (a twelfth – now of a foot). The pound (from *pondere* 'to weigh') is used as a translation of *libra* – balance – which is the origin of its abbreviation lb. Even the habit of counting in dozens instead of tens is Roman in origin. These and all other similarly funny units – like the system in which all units start with 'f', and which uses furlong/fortnight as its unit of velocity – are now officially defined as multiples of SI units.
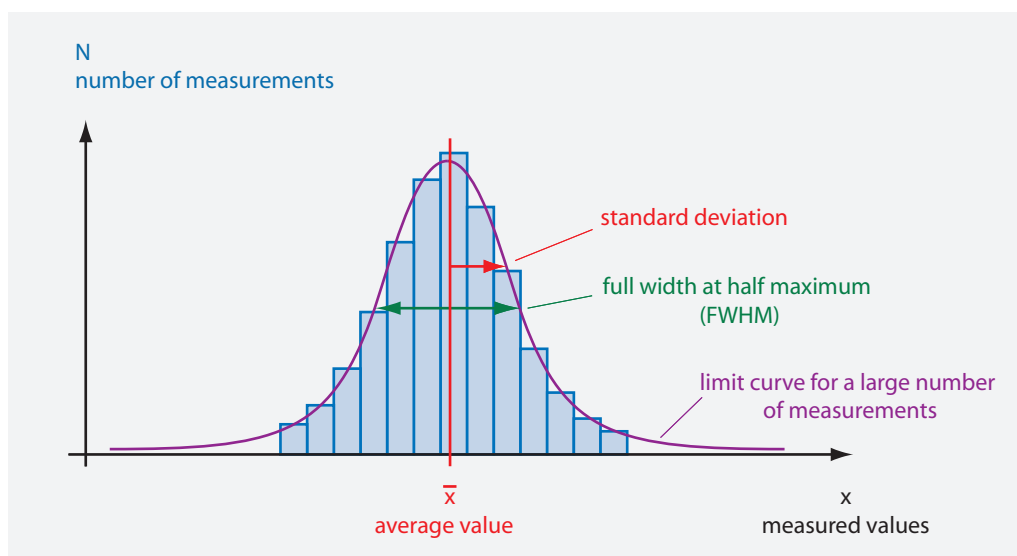
FIGURE 118 A precision experiment and its measurement distribution.

The tables below give the values of the most important physical constants and particle properties in SI units and in a few other common units, as published in the standard references. The values are the world averages of the best measurements made up to the present. As usual, experimental errors, including both random and estimated systematic errors, are expressed by giving the standard deviation in the last digits; e.g. 0.31(6) means – roughly speaking – 0.31 ± 0.06. In fact, behind each of the numbers in the following tables there is a long story which is worth telling, but for which there is not enough room here.

LIMITS TO PRECISION

What are the limits to accuracy and precision? There is no way, even in principle, to measure a length x to a *precision* higher than about 61 digits, because the ratio between the largest and the smallest measurable length is $\Delta x/x > l_{\mathrm{Pl}}/d_{\mathrm{horizon}} = 10^{-61}$. (Is this ratio valid also for force or for volume?) In the final volume of our text, studies of clocks and metre bars strengthen this theoretical limit.

But it is not difficult to deduce more stringent practical limits. No imaginable machine can measure quantities with a higher precision than measuring the diameter of the Earth within the smallest length ever measured, about 10^{-19} m; that is about 26 digits of precision. Using a more realistic limit of a 1000 m sized machine implies a limit of 22 digits. If, as predicted above, time measurements really achieve 17 digits of precision, then they are nearing the practical limit, because apart from size, there is an additional practical restriction: cost. Indeed, an additional digit in measurement precision often means an additional digit in equipment cost.

Physical constants

In principle, all quantitative properties of matter can be calculated with quantum theory. For example, colour, density and elastic properties can be predicted using the values of the following constants using the equations of the standard model of high-energy
physics.

TABLE 25 Basic physical constants.

QUANTITY	SYMBOL	VALUE IN SI UNITS	UNCERT.[a]
number of space-time dimensions		$3 + 1$	0 [b]
vacuum speed of light[c]	c	$299\,792\,458$ m/s	0
vacuum permeability[c]	μ_0	$4\pi \cdot 10^{-7}$ H/m	0
		$= 1.256\,637\,061\,435 \ldots$ µH/m	0
vacuum permittivity[c]	$\varepsilon_0 = 1/\mu_0 c^2$	$8.854\,187\,817\,620 \ldots$ pF/m	0
original Planck constant	h	$6.626\,068\,76(52) \cdot 10^{-34}$ Js	$7.8 \cdot 10^{-8}$
reduced Planck constant	$\hbar$	$1.054\,571\,596(82) \cdot 10^{-34}$ Js	$7.8 \cdot 10^{-8}$
positron charge	e	$0.160\,217\,646\,2(63)$ aC	$3.9 \cdot 10^{-8}$
Boltzmann constant	k	$1.380\,650\,3(24) \cdot 10^{-23}$ J/K	$1.7 \cdot 10^{-6}$
gravitational constant	G	$6.673(10) \cdot 10^{-11}$ Nm^2/kg^2	$1.5 \cdot 10^{-3}$
gravitational coupling constant	$\kappa = 8\pi G/c^4$	$2.076(3) \cdot 10^{-43}$ s^2/kg m	$1.5 \cdot 10^{-3}$
fine structure constant,[d]	$\alpha = \frac{e^2}{4\pi\varepsilon_0 \hbar c}$	$1/137.035\,999\,76(50)$	$3.7 \cdot 10^{-9}$
e.m. coupling constant	$= \alpha_{em}(m_e^2 c^2)$	$= 0.007\,297\,352\,533(27)$	$3.7 \cdot 10^{-9}$
Fermi coupling constant,[d]	$G_F/(\hbar c)^3$	$1.166\,39(1) \cdot 10^{-5}$ GeV^{-2}	$8.6 \cdot 10^{-6}$
weak coupling constant	$\alpha_w(M_Z) = g_w^2/4\pi$	$1/30.1(3)$	$1 \cdot 10^{-2}$
weak mixing angle	$\sin^2\theta_W(\overline{MS})$	$0.231\,24(24)$	$1.0 \cdot 10^{-3}$
weak mixing angle	$\sin^2\theta_W$ (on shell)	$0.2224(19)$	$8.7 \cdot 10^{-3}$
	$= 1 - (m_W/m_Z)^2$		
strong coupling constant[d]	$\alpha_s(M_Z) = g_s^2/4\pi$	$0.118(3)$	$25 \cdot 10^{-3}$

a. Uncertainty: standard deviation of measurement errors.
b. Only down to 10^{-19} m and up to 10^{26} m.
c. Defining constant.
d. All coupling constants depend on the 4-momentum transfer, as explained in the section on
renormalization. *Fine structure constant* is the traditional name for the electromagnetic coupling constant α in the case of a 4-momentum transfer of $Q^2 = m_e^2 c^2$, which is the smallest one possible. At higher momentum transfers it has larger values, e.g., $\alpha_{em}(Q^2 = M_W^2 c^2) \approx 1/128$. In contrast, the strong coupling constant has lover values at higher momentum transfers; e.g., $\alpha_s(34\,\text{GeV}) = 0.14(2)$.

Why do all these constants have the values they have? For any constant *with a dimension*, such as the quantum of action $\hbar$, the numerical value has only historical meaning. It is $1.054 \cdot 10^{-34}$ Js because of the SI definition of the joule and the second. The question why the value of a dimensional constant is not larger or smaller therefore always requires one to understand the origin of some dimensionless number giving the ratio between the constant and the corresponding natural unit that is defined with c, G, $\hbar$ and α.

Understanding the sizes of atoms, people, trees and stars, the duration of molecular and atomic processes, or the mass of nuclei and mountains, implies understanding the ratios between these values and the corresponding natural units. The key to understanding nature is thus the understanding of all ratios, and thus of all dimensionless constants. The quest of understanding all ratios, all dimensionless constants, including the fine structure constant α itself, is completed only in the final volume of our adventure.

The basic constants yield the following useful high-precision observations.

TABLE 26 Derived physical constants.

QUANTITY	SYMBOL	VALUE IN SI UNITS	UNCERT.
Vacuum wave resistance	$Z_0 = \sqrt{\mu_0/\varepsilon_0}$	376.730 313 461 77... Ω	0
Avogadro's number	N_A	$6.022\,141\,99(47) \cdot 10^{23}$	$7.9 \cdot 10^{-8}$
Rydberg constant [a]	$R_\infty = m_e c \alpha^2/2h$	10 973 731.568 549(83) m^{-1}	$7.6 \cdot 10^{-12}$
conductance quantum	$G_0 = 2e^2/h$	77.480 916 96(28) μS	$3.7 \cdot 10^{-9}$
magnetic flux quantum	$\varphi_0 = h/2e$	2.067 833 636(81) pWb	$3.9 \cdot 10^{-8}$
Josephson frequency ratio	$2e/h$	483.597 898(19) THz/V	$3.9 \cdot 10^{-8}$
von Klitzing constant	$h/e^2 = \mu_0 c/2\alpha$	25 812.807 572(95) Ω	$3.7 \cdot 10^{-9}$
Bohr magneton	$\mu_B = e\hbar/2m_e$	9.274 008 99(37) yJ/T	$4.0 \cdot 10^{-8}$
cyclotron frequency of the electron	$f_c/B = e/2\pi m_e$	27.992 4925(11) GHz/T	$4.0 \cdot 10^{-8}$
classical electron radius	$r_e = e^2/4\pi\varepsilon_0 m_e c^2$	2.817 940 285(31) fm	$1.1 \cdot 10^{-8}$
Compton wavelength of the electron	$\lambda_c = h/m_e c$	2.426 310 215(18) pm	$7.3 \cdot 10^{-9}$
	$\lambdabar_c = \hbar/m_e c = r_e/\alpha$	0.386 159 264 2(28) pm	$7.3 \cdot 10^{-9}$
Bohr radius [a]	$a_\infty = r_e/\alpha^2$	52.917 720 83(19) pm	$3.7 \cdot 10^{-9}$
nuclear magneton	$\mu_N = e\hbar/2m_p$	$5.050\,783\,17(20) \cdot 10^{-27}$ J/T	$4.0 \cdot 10^{-8}$
proton–electron mass ratio	m_p/m_e	1 836.152 667 5(39)	$2.1 \cdot 10^{-9}$
Stefan–Boltzmann constant	$\sigma = \pi^2 k^4/60\hbar^3 c^2$	56.704 00(40) nW/m^{2}K^4	$7.0 \cdot 10^{-6}$
Wien's displacement constant	$b = \lambda_{max}T$	2.897 768 6(51) mmK	$1.7 \cdot 10^{-6}$
bits to entropy conversion const.		10^{23} bit = 0.956 994 5(17) J/K	$1.7 \cdot 10^{-6}$
TNT energy content		3.7 to 4.0 MJ/kg	$4 \cdot 10^{-2}$

a. For infinite mass of the nucleus.

Some useful properties of our local environment are given in the following table.

TABLE 27 Astronomical constants.

QUANTITY	SYMBOL	VALUE
tropical year 1900 [a]	a	31 556 925.974 7 s
tropical year 1994	a	31 556 925.2 s
mean sidereal day	d	$23^h 56' 4.090\,53''$
astronomical unit [b]	AU	149 597 870.691(30) km
light year	al	9.460 528 173 ... Pm

TABLE 27 (Continued) Astronomical constants.

QUANTITY	SYMBOL	VALUE
parsec	pc	30.856 775 806 Pm = 3.261 634 al
Earth's mass	$M_⊕$	$5.973(1) \cdot 10^{24}$ kg
Geocentric gravitational constant	GM	$3.986\,004\,418(8) \cdot 10^{14}$ m^3/s^2
Earth's gravitational length	$l_⊕ = 2GM/c^2$	8.870 056 078(16) mm
Earth's equatorial radius [c]	$R_{⊕eq}$	6378.1366(1) km
Earth's polar radius [c]	$R_{⊕p}$	6356.752(1) km
Equator–pole distance [c]		10 001.966 km (average)
Earth's flattening [c]	$e_⊕$	1/298.25642(1)
Earth's av. density	$\rho_⊕$	5.5 Mg/m^3
Earth's age	$T_⊕$	4.50(4) Ga = 142(2) Ps
Moon's radius	$R_{☾ v}$	1738 km in direction of Earth
Moon's radius	$R_{☾ h}$	1737.4 km in other two directions
Moon's mass	$M_☾$	$7.35 \cdot 10^{22}$ kg
Moon's mean distance [d]	$d_☾$	384 401 km
Moon's distance at perigee [d]		typically 363 Mm, historical minimum 359 861 km
Moon's distance at apogee [d]		typically 404 Mm, historical maximum 406 720 km
Moon's angular size [e]		average 0.5181° = 31.08′, minimum 0.49°, maximum - shortens line 0.55°
Moon's average density	$\rho_☾$	3.3 Mg/m^3
Jupiter's mass	$M_♃$	$1.90 \cdot 10^{27}$ kg
Jupiter's radius, equatorial	$R_♃$	71.398 Mm
Jupiter's radius, polar	$R_♃$	67.1(1) Mm
Jupiter's average distance from Sun	$D_♃$	778 412 020 km
Sun's mass	$M_⊙$	$1.988\,43(3) \cdot 10^{30}$ kg
Sun's gravitational length	$l_⊙ = 2GM_⊙/c^2$	2.953 250 08 km
Sun's luminosity	$L_⊙$	384.6 YW
Solar equatorial radius	$R_⊙$	695.98(7) Mm
Sun's angular size		0.53° average; minimum on fourth of July (aphelion) 1888″, maximum on fourth of January (perihelion) 1952″
Sun's average density	$\rho_⊙$	1.4 Mg/m^3
Sun's average distance	AU	149 597 870.691(30) km
Sun's age	$T_⊙$	4.6 Ga
Solar velocity around centre of galaxy	$v_{⊙g}$	220(20) km/s
Solar velocity against cosmic background	$v_{⊙b}$	370.6(5) km/s
Distance to Milky Way's centre		8.0(5) kpc = 26.1(1.6) kal

TABLE 27 (Continued) Astronomical constants.

QUANTITY	SYMBOL	VALUE
Milky Way's age		13.6 Ga
Milky Way's size		$c.\ 10^{21}$ m or 100 kal
Milky Way's mass		10^{12} solar masses, $c.\ 2 \cdot 10^{42}$ kg
Most distant galaxy cluster known	SXDF-XCLJ 0218-0510	$9.6 \cdot 10^{9}$ al

a. Defining constant, from vernal equinox to vernal equinox; it was once used to define the second. (Remember: π seconds is about a nanocentury.) The value for 1990 is about 0.7 s less, corresponding to a slowdown of roughly 0.2 ms/a. (Watch out: why?) There is even an empirical formula for the change of the length of the year over time.

b. Average distance Earth–Sun. The truly amazing precision of 30 m results from time averages of signals sent from Viking orbiters and Mars landers taken over a period of over twenty years.

c. The shape of the Earth is described most precisely with the World Geodetic System. The last edition dates from 1984. For an extensive presentation of its background and its details, see the www.wgs84.com website. The International Geodesic Union refined the data in 2000. The radii and the flattening given here are those for the 'mean tide system'. They differ from those of the 'zero tide system' and other systems by about 0.7 m. The details constitute a science in itself.

d. Measured centre to centre. To find the precise position of the Moon at a given date, see the www.fourmilab.ch/earthview/moon_ap_per.html page. For the planets, see the page www.fourmilab.ch/solar/solar.html and the other pages on the same site.

e. Angles are defined as follows: 1 degree = $1° = π/180$ rad, 1 (first) minute = $1' = 1°/60$, 1 second (minute) = $1'' = 1'/60$. The ancient units 'third minute' and 'fourth minute', each 1/60th of the preceding, are not in use any more. ('Minute' originally means 'very small', as it still does in modern English.)

Some properties of nature at large are listed in the following table. (If you want a challenge, can you determine whether any property of the universe itself is listed?)

TABLE 28 Astrophysical constants.

QUANTITY	SYMBOL	VALUE
gravitational constant	G	$6.672\ 59(85) \cdot 10^{-11}$ m^3/kg s^2
cosmological constant	Λ	$c.\ 1 \cdot 10^{-52}$ m^{-2}
age of the universe [a]	t_0	$4.333(53) \cdot 10^{17}$ s $= 13.73(0.17) \cdot 10^{9}$ a
(determined from space-time, via expansion, using general relativity)		
age of the universe [a]	t_0	over $3.5(4) \cdot 10^{17}$ s $= 11.5(1.5) \cdot 10^{9}$ a
(determined from matter, via galaxies and stars, using quantum theory)		
Hubble parameter [a]	H_0	$2.3(2) \cdot 10^{-18}$ s^{-1} $= 0.73(4) \cdot 10^{-10}$ a^{-1}
		$= h_0 \cdot 100$ km/s Mpc $= h_0 \cdot 1.0227 \cdot 10^{-10}$ a^{-1}
reduced Hubble parameter [a]	h_0	$0.71(4)$
deceleration parameter	$q_0 = -(\ddot{a}/a)_0/H_0^2$	$-0.66(10)$
universe's horizon distance [a]	$d_0 = 3ct_0$	$40.0(6) \cdot 10^{26}$ m $= 13.0(2)$ Gpc
universe's topology		trivial up to 10^{26} m
number of space dimensions		3, for distances up to 10^{26} m

TABLE 28 (Continued) Astrophysical constants.

QUANTITY	SYMBOL	VALUE
critical density of the universe	$\rho_c = 3H_0^2/8\pi G$	$h_0^2 \cdot 1.878\,82(24) \cdot 10^{-26}\,\text{kg/m}^3$ $= 0.95(12) \cdot 10^{-26}\,\text{kg/m}^3$
(total) density parameter [a]	$\Omega_0 = \rho_0/\rho_c$	1.02(2)
baryon density parameter [a]	$\Omega_{B0} = \rho_{B0}/\rho_c$	0.044(4)
cold dark matter density parameter [a]	$\Omega_{CDM0} = \rho_{CDM0}/\rho_c$	0.23(4)
neutrino density parameter [a]	$\Omega_{v0} = \rho_{v0}/\rho_c$	0.001 to 0.05
dark energy density parameter [a]	$\Omega_{X0} = \rho_{X0}/\rho_c$	0.73(4)
dark energy state parameter	$w = p_X/\rho_X$	$-1.0(2)$
baryon mass	m_b	$1.67 \cdot 10^{-27}\,\text{kg}$
baryon number density		$0.25(1)\,/\text{m}^3$
luminous matter density		$3.8(2) \cdot 10^{-28}\,\text{kg/m}^3$
stars in the universe	n_s	$10^{22\pm1}$
baryons in the universe	n_b	$10^{81\pm1}$
microwave background temperature [b]	T_0	2.725(1) K
photons in the universe	n_γ	10^{89}
photon energy density	$\rho_\gamma = \pi^2 k^4/15 T_0^4$	$4.6 \cdot 10^{-31}\,\text{kg/m}^3$
photon number density		$410.89\,/\text{cm}^3$ or $400\,/\text{cm}^3 (T_0/2.7\,\text{K})^3$
density perturbation amplitude	$\sqrt{S}$	$5.6(1.5) \cdot 10^{-6}$
gravity wave amplitude	$\sqrt{T}$	$< 0.71\sqrt{S}$
mass fluctuations on 8 Mpc	σ_8	0.84(4)
scalar index	n	0.93(3)
running of scalar index	$dn/d\ln k$	-0.03(2)
Planck length	$l_{Pl} = \sqrt{\hbar G/c^3}$	$1.62 \cdot 10^{-35}\,\text{m}$
Planck time	$t_{Pl} = \sqrt{\hbar G/c^5}$	$5.39 \cdot 10^{-44}\,\text{s}$
Planck mass	$m_{Pl} = \sqrt{\hbar c/G}$	$21.8\,\mu\text{g}$
instants in history [a]	t_0/t_{Pl}	$8.7(2.8) \cdot 10^{60}$
space-time points inside the horizon [a]	$N_0 = (R_0/l_{Pl})^3 \cdot$ (t_0/t_{Pl})	$10^{244\pm1}$
mass inside horizon	M	$10^{54\pm1}\,\text{kg}$

a. The index 0 indicates present-day values.
b. The radiation originated when the universe was 380 000 years old and had a temperature of about 3000 K; the fluctuations ΔT_0 which led to galaxy formation are today about $16 \pm 4\,\mu\text{K} = 6(2) \cdot 10^{-6}\,T_0$.

USEFUL NUMBERS

π	3.14159 26535 89793 23846 26433 83279 50288 41971 69399 37510$_5$
e	2.71828 18284 59045 23536 02874 71352 66249 77572 47093 69995$_9$
γ	0.57721 56649 01532 86060 65120 90082 40243 10421 59335 93992$_3$
$\ln 2$	0.69314 71805 59945 30941 72321 21458 17656 80755 00134 36025$_5$
$\ln 10$	2.30258 50929 94045 68401 79914 54684 36420 76011 01488 62877$_2$
$\sqrt{10}$	3.16227 76601 68379 33199 88935 44432 71853 37195 55139 32521$_6$
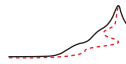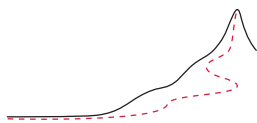
CHALLENGE HINTS AND SOLUTIONS

Challenge 1, page 9: Do not hesitate to be demanding and strict. The next edition of the text will benefit from it.

Challenge 4, page 19: The liquid drops have to detach from the flow exactly inside the metal counter-electrodes. There is always a tiny charge somewhere on the metal structures (due to cosmic rays, rubbing, previous charging, etc.). In Figure 119, this initial charge is the positive charge drawn on the lower left and upper right metal structure. When the water droplets form, they get a charge that is opposite to that of the metal surrounding the region where they form. The negative droplets fall into the other metal structure. Through the negative charge accumulating there, the positive charge in the first structure increases. When the charge on the metal structure increases, the charge separation in the droplets is more efficient. In other words, water droplet formation inside the metal structures *amplifies* any initial charge. After a while, the charge value and the associated voltage are so high that it leads to a loud bang (if everything is dry, including the air.) Then the process starts again. In fact, a vaguely similar charge separation mechanism leads to cloud charging and to lightning. If you want to build a Kelvin generator at home, have a look at the de.wikipedia.org/wiki/Kelvin-Generator web page, or search for other internet site on the topic.

To avoid the sparks in the fuel tanks of its cars, Opel simply earthed the metal tube at the tank inlet; they had forgotten to ensure electric contact between the tube and the rest of the car.

The explosion of fuel can also occur if you pour fuel into your car from a metal container. Several times, fuel thieves were 'punished' by an explosion triggered by static electricity when they tried to pour stolen fuel into their own car.

On every airport you can see how the problem is avoided: before even attaching the fuel tube to an aeroplane, the worker attaches a conducting cable to connect the truck (or the tank) to the aeroplane.

Challenge 5, page 20: We look at the two spark through a rapidly rotating mirror. In this way, small timing differences lead to position differences of the two sparks. In the 19th century, the speed values measured in this way varied between 6000 km/s and way over 100 000 km/s, because the speed depends on the effective capacitance and inductance of wire and set-up. Only if these effects can be neglected is the measured speed the same as that of light in vacuum, namely around 300 000 km/s. In modern cables, the speed is often around a third of this value.

Challenge 6, page 21: A lot of noise appeared while the metal pendulum banged wildly between the two fixed bells.

Challenge 8, page 24: No.

Challenge 9, page 25: The field at a distance of 1 m from an electron is 1.4 nV/m.

Challenge 10, page 25: A simple geometrical effect: anything flowing out homogeneously from a sphere diminishes with the square of the distance.

Challenge 11, page 27: One has $F = \alpha \hbar c N_A^2 / 4R^2 = 3 \cdot 10^{12}$ N, an enormous force, corresponding to the weight of 300 million tons. It shows the enormous forces that keep matter together.
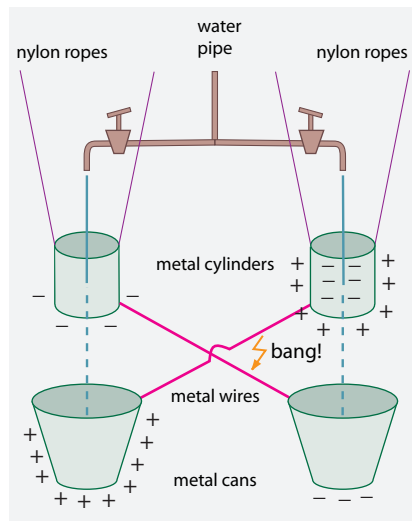
FIGURE 119 The key process in the Kelvin generator: charge separation during droplet formation.

Obviously, there is no way to keep 1 g of positive charge together, as the repulsive forces among the charges would be even larger.

Challenge 12, page 27: To show the full equivalence of Coulomb's and Gauss's 'laws', first show that it holds for a single point charge. Then expand the result for more than one point charge. That gives Gauss's 'law' in integral form, as given just before this challenge.

To deduce the integral form of Gauss's 'law' for a single point charge, one has to integrate over the closed surface. The essential point here is to note that the integration can be carried out for an inverse square dependence only. This dependence allows transforming the scalar product between the local field and the area element into a normal product between the charge and the solid angle Ω:

$$E \, \mathrm{d}A = \frac{q \mathrm{d}A \cos\theta}{4\pi\varepsilon_0 r^2} = \frac{q \mathrm{d}\Omega}{4\pi\varepsilon_0} \; . \tag{104}$$

In case that the surface is closed the integration is then straightforward.

To deduce the differential form of (the static) Gauss's 'law', namely

$$\nabla E = \frac{\rho}{\varepsilon_0} \; , \tag{105}$$

make use of the definition of the charge density ρ and of the purely mathematical relation

$$\oint_{\text{closed surface}} E \, \mathrm{d}A = \int_{\text{enclosed volume}} \nabla E \, \mathrm{d}V \; , \tag{106}$$

This mathematical relation, valid for any vector field E, is called *Gauss's theorem*. It simply states that the flux is the volume integral of the divergence.

To deduce the full form of Gauss's law, including the time-derivative of the magnetic field, include relativistic effects by changing viewpoint to a moving observer.

Challenge 14, page 27: Uncharged bodies can attract each other if they are made of charged constituents neutralizing each other, and if the charges are constrained in their mobility. The charge fluctuations then lead to attraction. Most molecules interact among each other in this way; such forces are also at the basis of surface tension in liquids and thus of droplet formation.

Challenge 15, page 27: No; batteries only separate charges and pump them around.

Challenge 17, page 30: The ratio q/m of electrons and that of the free charges inside metals is not exactly the same.

Challenge 19, page 32: Find out a way to test the issue, perform the experiment, and publish it!

Challenge 20, page 37: If you can, publish the result. Researchers have tried to put people on the ocean during clouded days, have tried experiments in dark rooms, but nothing has been found so far. The experiences of people in magnetic resonance imaging equipment is inconclusive so far.

Challenge 22, page 42: No.

Challenge 24, page 43: The correct version of Ampère's 'law' is

$$\nabla \times \boldsymbol{B} - \frac{1}{c^2}\frac{\partial \boldsymbol{E}}{\partial t} = \mu_0 \boldsymbol{j} \tag{107}$$

whereas the expression mentioned in the text misses the term $\frac{\partial \boldsymbol{E}}{\partial t}$.

For another way to state the difference, see RICHARD P. FEYNMAN, ROBERT B. LEIGHTON & MATTHEW SANDS, *The Feynman Lectures on Physics*, volume II, Addison Wesley, p. 21-1, 1977.

Challenge 25, page 43: Only boosts with relativistic speeds mix magnetic and electric fields to an appreciable amount.

Challenge 27, page 45: The dual field $*F$ is defined on page 65.

Challenge 28, page 45: Scalar products of four vectors are always, by construction, Lorentz invariant quantities.

Challenge 32, page 47: Almost all neutral particles are made of charged ones. So the argument holds for them as well. There is only one exception: neutrinos. For them, the argument is not valid.

Challenge 33, page 47: The signal speed in metals is higher than the electron speed, namely around the speed of light. This speed determines the magnetic field strength.

Challenge 34, page 48: The illumination of the sun changes the ionization in the upper atmosphere and provokes convection in the ionosphere. The tides move the ions in the ocean and in
the atmosphere. These currents lead to magnetic fields which can be seen in sensitive compass needles.

Challenge 35, page 48: If you find such an effect and are able to demonstrate it, publish it in a didactic journal.

Challenge 36, page 48: Usually, the cables of high voltage lines are too warm to be comfortable.

Challenge 37, page 49: Move them to form a T shape.

Challenge 38, page 49: Hint: a shining bulb is hot.

Challenge 39, page 49: For three and more switches, on uses inverters; an inverter is a switch with two inputs and two outputs which in one position, connects first and second input to first and second output respectively, and in the other position connects the first input to the second output and vice versa. (There are other possibilities, though; wires can be saved using electromagnetic relay switches.)

Challenge 41, page 50: Blond children tend to have the thinnest hair, thus giving the greatest effect. Dry weather is needed to avoid that the moisture in the air discharges the head thus preventing the hair to raise at all.

Challenge 42, page 50: It is possible; however, the systems so far are not small and are dangerous for human health. The idea to collect solar power in deep space and then beam it to the Earth as microwaves has often been aired. Finances and dangers have blocked it so far.

Challenge 44, page 51: Glue two mirrors together at a right angle. Or watch yourself on TV using a video camera.

Challenge 45, page 51: This is again an example of combined triboluminescence and triboelectricity. See also the websites scienceworld.wolfram.com/physics/Triboluminescence.html and www.geocities.com/RainForest/9911/tribo.htm.

Challenge 48, page 53: Pepper is lighter than salt, and thus reacts to the spoon before the salt does.

Challenge 49, page 54: For a wavelength of 546.1 nm (standard green), that is a bit over 18 wavelengths.

Challenge 50, page 54: The angular size of the Sun is too large; diffraction plays no role here.

Challenge 51, page 54: Just use a high speed camera.

Challenge 52, page 54: The current flows perpendicularly to the magnetic field and is thus deflected. It pulls the whole magnet with it.

Challenge 53, page 55: The most simple equivalent to a coil is a rotating mass being put into rotation by the flowing water. A transformer would then be made of two such masses connected through their axis.

Challenge 54, page 56: Light makes seven turns of the Earth in one second.

Challenge 58, page 56: There are no permanent magnets in nature that fit in a floor and that are strong enough to achieve a floating height of 50 to 80 cm. (Note that in one image the floating height is so large that the legs of the woman do not touch the floor.) And anybody who has tried this with an electromagnet knows that such a device would be larger than a complete room.

Looking carefully at the images, you will also note that they are not photographs: there are errors with the shadow and with the reflected image of the woman. And most of all, nobody would cut half the bed out of an image with a woman on the bed. Finally, nobody has ever seen the floating bed shown in the images.

Challenge 61, page 59: The charged layer has the effect that almost only ions of one charge pass the channels. As a result, charges are separated on the two sides of the liquid, and a current is generated.

Challenge 62, page 59: Leakage currents change the picture. The long term voltage ratio is given by the leakage resistance ratio $V_1/V_2 = R_1/R_2$, as can be easily verified in experiments.

Challenge 70, page 68: Some momentum is carried away by the electromagnetic field.

Challenge 71, page 69: Field lines and equipotential surfaces are always orthogonal to each other. Thus a field line cannot cross an equipotential surface twice.

Challenge 83, page 76: The argument is the same as for the increase in entropy: reverse processes are possible, but the probability is so low that they do not appear in practice. The extremely low probability is due to the fluctuations induced by the environment.

Challenge 84, page 76: Just draw a current through a coil with its magnetic field, then draw the mirror image of the current and redraw the magnetic field.

Challenge 85, page 77: Other asymmetries in nature include the helicity of the DNA molecules making up the chromosomes and many other molecules in living systems, the right hand preference of most humans, the asymmetry of fish species which usually stay flat on the bottom of the seas.

Challenge 86, page 77: This is not possible at all using gravitational or electromagnetic systems or effects. The only way is to use the weak nuclear interaction, as shown in the chapter on the nucleus.

Challenge 87, page 77: The Lagrangian does not change if one of the three coordinates is changed by its negative value.

Challenge 88, page 77: The image flips up: a 90 degree rotation turns the image by 180 degrees.

Challenge 89, page 78: Imagine E and B as the unite vectors of two axes in complex space. Then any rotation of these axes is also a generalized duality symmetry.

Challenge 90, page 81: The angular momentum was put into the system when it was formed. If we bring a point charge from infinity along a straight line to its final position close to a magnetic dipole, the magnetic force acting on the charge is not directed along the line of motion. It therefore creates a non-vanishing torque about the origin. See J. M. AGUIRREGABIRIA & A. HERNANDEZ, The Feynman paradox revisited, *European Journal of Physics* 2, pp. 168–170, 1981.

Challenge 91, page 81: Show that even though the radial magnetic field of a spherical wave is vanishing by definition, Maxwell's equations would require it to be different from zero. Since electromagnetic waves are transversal, it is also sufficient to show that it is impossible to comb a hairy sphere without having a (double) vortex or two simple vortices. Despite these statements, quantum theory changes the picture somewhat: the emission probability of a photon from an excited atom in a degenerate state is spherically symmetric exactly.

Challenge 95, page 81: The emitted radiation is strongly suppressed because the size of the dipole (the plug) is much smaller than the wavelength of the field.

Challenge 97, page 87: In every case of interference, the energy is redistributed into other directions. This is the general rule; sometimes it is quite tricky to discover this other direction.

Challenge 98, page 87: The author regularly sees about 7 lines; assuming that the distance is around 20 μm, this makes about 3 μm per line. The wavelength must be smaller than this value and the frequency thus larger than 100 THz. The actual values for various colours are given in the table of the electromagnetic spectrum.

Challenge 100, page 88: He noted that when a prism produces a rainbow, a thermometer placed in the region after the colour red shows a temperature rise.

Challenge 101, page 90: Birefringence appears when the refraction is polarization dependent. Only two linear independent polarizations are possible, thus there is no trirefringence in nature. This holds true also for crystals which have three different indices of refraction in three directions!

Challenge 102, page 90: Light reflected form a water surface is partly polarized. Mirages are not.

Challenge 105, page 97: Figure 120 shows electrical field lines. We assume that the charge moves at a initial velocity v_0 that is small compared to c and that it decelerates to zero velocity during a time t_0. After a time T has elapsed, the radiation pulse has travelled a distance $R = cT$, where $T \gg t_0$. The figure shows that at a given kink, drawn in red, the ratio of the transverse field E_t and of the radial field E_r is given by the steepness of the of the kink. (Why?) Geometry then leads to

$$\frac{E_t}{E_r} = \frac{v_0 T \sin\theta}{c t_0} = \frac{aR \sin\theta}{c^2} \ . \tag{108}$$

Inserting Coulomb's expression for the radial field we get

$$E_t = \frac{1}{4\pi\varepsilon_0 c^2} \frac{a \sin\theta}{R} \ . \tag{109}$$
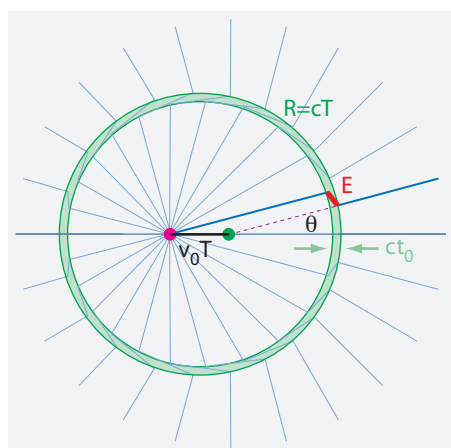
FIGURE 120 Calculating the transverse field of an accelerated charge.

The magnitude of the transversal field thus decreases with $1/R$. In addition, the field depends on the angle θ; this is clearly visible both in Figure 120 and in Figure 56 on page 98. In other words, transmitter antennas have a preferred direction of power emission, namely perpendicularly to the direction of acceleration.

Challenge 106, page 99: Such an observer would experience a wavy but static field, which cannot exist, as the equations for the electromagnetic field show.

Challenge 107, page 100: You would never die. Could you reach the end of the universe?

Challenge 110, page 101: A surface of $1\,\mathrm{m}^2$ perpendicular to the light receives about $1\,\mathrm{kW}$ of radiation. It generates the same pressure as the weight of about 0.3 mg of matter. That generates $3\,\mu\mathrm{Pa}$ for black surfaces, and the double for mirrors.

Challenge 112, page 102: The shine side gets twice the momentum transfer as the black side, and thus should be pushed backwards.

Challenge 115, page 103: A polarizer can do this.

Challenge 118, page 104: The interference patterns change when colours are changed. Rainbows also appear because different colours are due to different frequencies.

Challenge 120, page 105: The full rainbow is round like a circle. You can produce one with a garden hose, if you keep the hose in your hand while you stand on a chair, with your back to the evening Sun. (Well, one small part is missing; can you imagine which part?) The circle is due to the spherical shape of droplets. If the droplets were of different shape, *and* if they were all aligned, the rainbow would have a different shape than a simple circle.

Challenge 125, page 108: Take a film of a distant supernova explosion, or better, an optical or gamma ray burst, and check whether it happens at the same time for each colour separately. This has been done extensively, and no differences have been detected within experimental errors.

Challenge 127, page 111: The first part of the forerunner is a feature with the shortest possible effective wavelength; thus it is given by taking the limit for infinite frequency.

Challenge 128, page 111: The light is pulsed; thus it is the energy velocity.

Challenge 129, page 111: Inside matter, the energy is transferred to atoms, then back to light, then to the next atoms, etc. That takes time and slows down the propagation.

Challenge 130, page 112: This is true even in general relativity, when the bending of the vacuum is studied.

Challenge 131, page 113: Almost no light passes; the intensity of the little light that is transmitted depends exponentially on the ratio between wavelength and hole diameter. One also says that after the hole there is an evanescent wave.

Challenge 134, page 118: The average temperature of the Earth is thus 287 K. The energy from the Sun is proportional to the fourth power of the temperature. The energy is spread (roughly) over half the Earth's surface. The same energy, at the Sun's surface, comes from a much smaller surface, given by the same angle as the Earth subtends there. We thus have $E \sim 2\pi R_{\text{Earth}}^2 T_{\text{Earth}}^4 = T_{\text{Sun}}^4 R_{\text{Earth}}^2 \alpha^2$, where α is half the angle subtended by the Sun. As a result, the temperature of the Sun is estimated to be $T_{\text{Sun}} = (T_{\text{Earth}}^4/\alpha^2)^{0.25} = 4 \text{ kK}$.

Challenge 138, page 119: Because the maximum of a spectrum in wavelengths and in frequencies is not the same, thus does *not* and cannot follow $c = f\lambda$.

Challenge 141, page 119: At high temperature, all bodies approach black bodies. The colour is more important than other colour effects. The oven and the objects have the same temperature. Thus they cannot be distinguished from each other. To do so nevertheless, illuminate the scene with powerful light and then take a picture with small sensitivity. Thus one always needs bright light to take pictures of what happens inside fires.

Challenge 145, page 123: For small mirrors or lenses, like those used in microscopes, mass production is easier for lenses. In contrast, large mirrors are much easier and cheaper to fabricate than large lenses, because mirrors use less glass, are lighter, and allow changing their shape with actuators.

Challenge 146, page 124: Syrup shows an even more beautiful effect in the following setting. Take a long transparent tube closed at one end and fill it with syrup. Shine a red helium–neon laser into the tube from the bottom. Then introduce a linear polarizer into the beam: the light seen in the tube will form a spiral. By rotating the polarizer you can make the spiral advance or retract. This effect, called the *optical activity* of sugar, is due to the ability of sugar to rotate light polarization and to a special property of plants: they make only one of the two mirror forms of sugar.

Challenge 147, page 126: The thin lens formula is

$$\frac{1}{d_{\text{o}}} + \frac{1}{d_{\text{i}}} = \frac{1}{f} \; . \tag{110}$$

It is valid for diverging and converging lenses, as long as their own thickness is negligible. The strength of a lens can thus be measured with the quantity $1/f$. The unit 1 m^{-1} is called a *diopter*; it is used especially for reading glasses. Converging lenses have positive, diverging lenses negative values.

However, the thin lens formula is only an approximation, and is never used in lens design. It is a relic of old textbooks. Modern lens designers always use Gaussian optic for calculations. (See, for example, FRANCIS A. JENKINS & HARVEY E. WHITE, *Fundamentals of Optics*, McGraw-Hill, 1957.)

Challenge 149, page 127: A light microscope is basically made of two converging lenses. One lens – or lens system – produces an enlarged real image and the second one produces an enlarged virtual image of the previous real image. Figure 121 also shows that microscopes always turn images upside down. Due to the wavelength of light, light microscopes have a maximum resolution of about 1 μm. Note that the magnification of microscopes is unlimited; what is limited is their resolution. This is exactly the same behaviour shown by digital images. The *resolution* is simply the size of the smallest possible pixel that makes sense.

To learn more about microscopes, read the beautiful text by ELIZABETH M. SLATER & HENRY S. SLATER, *Light and Electron Microscopy*, Cambridge University Press, 1993.
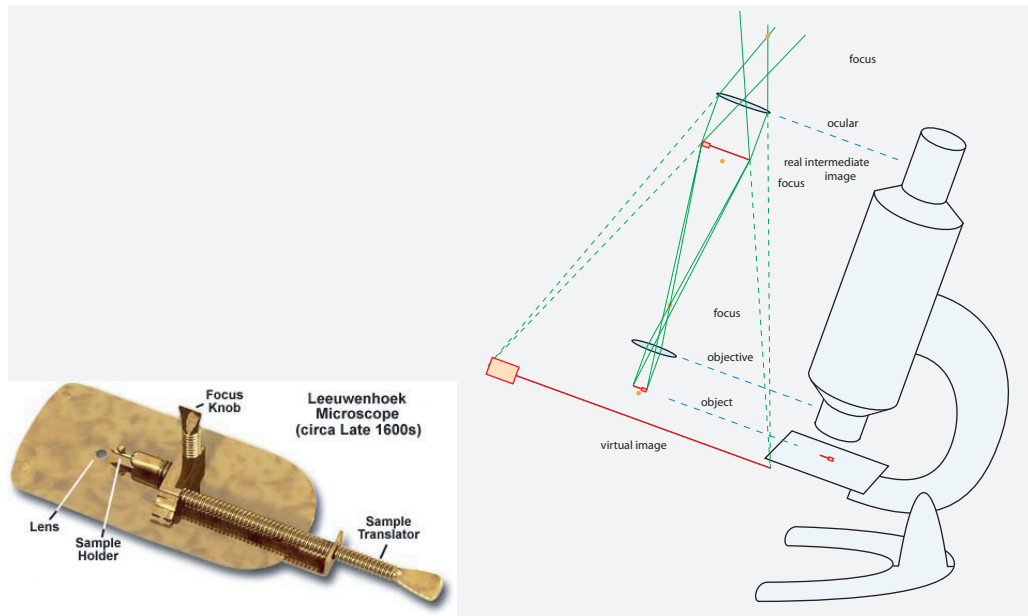
FIGURE 121 One lens made the oldest commercial microscope, from 1680 (length *c.* 8 cm, to be held close to the eye), but two converging lenses make a modern microscope (photo WikiCommons).
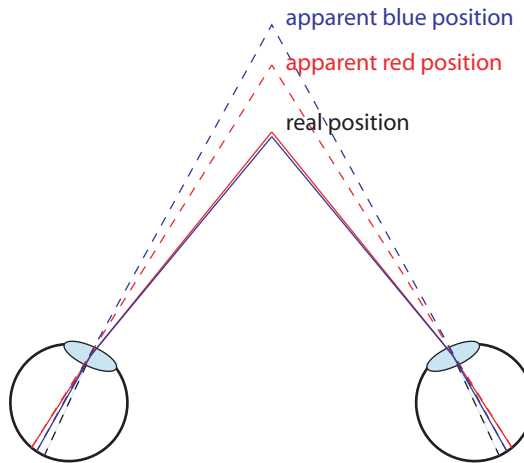
FIGURE 122 The relation between the colour depth effect and the lens dispersion of the human eye.

Challenge 151, page 129: The dispersion at the lens leads to different apparent image positions, as shown in Figure 122. For more details on the dispersion in the human eye and the ways of using

it to create three-dimensional effects, see the article by C. UCKE & R. WOLF, *Durch Farbe in die dritte Dimension*, *Physik in unserer Zeit* **30**, pp. 50–53, 1999.

Challenge 152, page 132: The 1 mm beam would return 1000 times as wide as the 1 m beam. A perfect 1 m-wide beam of green light would be 209 m wide on the Moon; can you deduce this result from the (important) formula that involves distance, wavelength, initial diameter and final diameter? Try to guess this beautiful formula first, and then deduce it. In reality, the values are a few times larger than the theoretical minimum thus calculated. See the www.csr.utexas.edu/mlrs and ilrs.gsfc.nasa.gov websites.

Challenge 153, page 132: It is often said that evolution tuned the number of cones in the eye to the maximum resolution with open pupil; the experts on the subject however maintain that there are somewhat larger numbers of cones.

Challenge 154, page 132: The answer should lie between one or two dozen kilometres, assuming ideal atmospheric circumstances.

Challenge 157, page 136: In fact, there is no way that a hologram of a person can walk around and frighten a real person. A hologram is always transparent; one can always see the background through the hologram. A hologram thus always gives an impression similar to what moving pictures usually show as ghosts.

Challenge 158, page 137: There is a blind spot in the eye; that is a region in which images are not perceived. The brain than assumes that the image at that place is the same than at its borders. If a spot falls exactly inside it, it disappears.

Challenge 160, page 140: The eye and brain surely do not switch the up and the down direction at a certain age. Besides, where does the idea come from that babies see upside-down?

Challenge 161, page 141: The eye and vision system subtract patterns that are constant in time.

Challenge 162, page 142: Not really; a Cat's-eye uses two reflections at the sides of a cube. A living cat's eye has a large number of reflections. The end effect is the same though: light returns back to the direction it came from.

Challenge 163, page 142: At 10 pc=32.6 al, the Sun would have apparent magnitude 4.7. At 20 pc=65.2 al, it would appear 4 times fainter, thus about 1.5 magnitudes more, therefore with an apparent visual magnitude of about 6.2. This is near the limit magnitude of the eye. The actual limiting magnitude of the eye is neither constant nor universal, so the distance of 50 light years is not a sharp limit. The limiting magnitude, – like the night vision, or *scotopic sensitvity* – depends on the partial pressure of oxygen in the atmosphere the observer is breathing, on the clarity of the air, on the zenith distance, and, above all, on the degree of dark adaptation. An eye exposed to the full brightness of the night sky in a very dark location far from light pollution is still not completely dark-adapted. You can easily see 7th-magnitude stars by blocking off most of the sky and just looking at a little patch of it. Some observers, under ideal conditions, have reliably reported seeing stars near 8th magnitude.

Challenge 164, page 142: The green surface seen at a low high angle is larger than when seen vertically, where the soil is also seen; the soil is covered by the green grass in low angle observation.

Challenge 165, page 142: It is indeed true. Modern telescopes have a large surface collecting light (up to 50 m^2) and have extremely sensitive detectors. The number of photons emitted by a match lit on the moon into the direction of a large telescope (how many?) is sufficient to trigger the detector.

Challenge 166, page 142: Of course not, as the group velocity is not limited by special relativity. The energy velocity is limited, but is not changed in this experiments.

Challenge 169, page 144: The Prussian explorer Alexander von Humboldt extensively checked this myth in the nineteenth century. He visited many mine pits and asked countless mine workers in Mexico, Peru and Siberia about their experiences. He also asked numerous chimney-sweeps. Neither him nor anybody else had ever seen the stars during the day.

Challenge 170, page 145: Watch the Sun with closed eyes, and remember the shade of red you see. Go into a closed room, turn a light bulb on, and watch it with closed eyes. Chose the distance from the bulb that yields the same shade of red. Then deduce the power of the Sun from the power of the light bulb and the inverse square dependence.

Challenge 172, page 148: The human body is slightly conducting and changes the shape of the field and thus effectively short circuits it. Usually, the field cannot be used to generate energy, as the currents involved are much too small. (Lightning bolts are a different story, of course. They are due – very indirectly – to the field of the Earth, but they are too irregular to be used consistently. Franklin's lightning rod is such an example.)

Challenge 173, page 150: The field at the surface of a sphere of radius r is given by $E = Q/4\pi\varepsilon_0 r^2$. Inserting $E = 200\,\text{V/m}$, one gets $Q = 0.9\,\text{MC}$.

Challenge 174, page 152: If you find a method that is different from the known estimates, publish it.

Challenge 180, page 157: All the illusions of the flying act look as if the magician is hanging on lines, as observed by many, including myself. (Photographic flashes are forbidden, a shimmery background is set up to render the observation of the lines difficult, no ring is ever actually pulled over the magician, the aquarium in which he floats is kept open to let the fishing lines pass through, always the same partner is 'randomly' chosen from the public, etc.) Information from eyewitnesses who have actually seen the fishing lines used by David Copperfield explains the reasons for these set-ups. The usenet news group alt.magic.secrets, in particular Tilman Hausherr, was central in clearing up this issue in all its details, including the name of the company that made the suspension mechanism.

Challenge 182, page 158: Any new one is worth a publication.

Challenge 183, page 161: Sound energy is also possible, as is mechanical work.

Challenge 184, page 163: Space-time deformation is not related to electricity; at least at everyday energies. Near Planck energies, this might be different, but nothing has been predicted yet.

Challenge 186, page 165: Ideal absorption is blackness (though it can be redness or whiteness at higher temperatures).

Challenge 187, page 165: Indeed, the Sun emits about $4 \cdot 10^{26}\,\text{W}$ from its mass of $2 \cdot 10^{30}\,\text{kg}$, about $0.2\,\text{mW/kg}$. The adult human body (at rest) emits about $100\,\text{W}$ (you can check this in bed at night), thus about $1.2\,\text{W/kg}$. This is about 6000 times more than the Sun.

Challenge 188, page 165: The charges on a metal box rearrange so that the field inside remains vanishing. This makes cars and aeroplanes safe against lightning. Of course, if the outside field varies so quickly that the rearrangement cannot follow, fields *can* enter the Faraday cage. (By the way, also fields with long wavelengths penetrate metals; specialized remote controls for opening security doors regularly use frequencies of $25\,\text{kHz}$ to achieve this.) However, one should wait a bit before stepping out of a car after lightning has hit, as the car is on rubber wheels with low conduction; waiting gives the charge time to flow into the ground.

For gravity and solid cages, mass rearrangement is not possible, so that there is no gravity shield.

Challenge 192, page 166: This is a touchy topic. It is not clear whether $50\,\text{Hz}$ fields are dangerous to humans. There is a high probability that they are not; but the issue is not settled.

Challenge 193, page 167: The number of photons times the quantum of action $\hbar$.
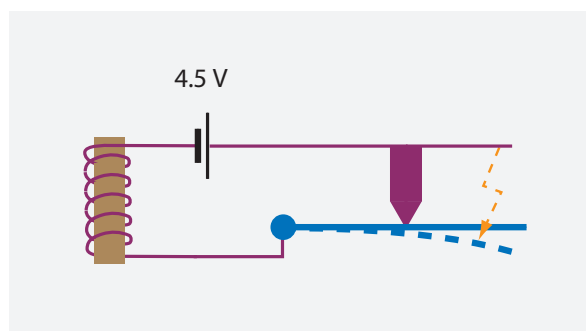
FIGURE 123 How to get electrical shocks from a 4.5 V pocket battery.

Challenge 195, page 167: The charging stops because a negatively charged satellite repels electrons and thus stops any electron collecting mechanism. Electrons are captured more frequently than ions because it is easier for them than for ions to have an inelastic collision with the satellite, due to their larger speed at a given temperature.

Challenge 196, page 167: Any loss mechanism will explain the loss of energy, such as electrical resistance or electromagnetic radiation. After a fraction of a second, the energy will be lost. This little problem is often discussed on the internet.

Challenge 197, page 167: Use the wire as shown in Figure 123. *If* the oscillation is properly tuned in frequency, and *if* the contact detaches properly at the tip, and *if* you touch the two contacts with a strong grip, you will get a stronger shock than you can stand.

Challenge 199, page 168: This should be possible in the near future; but both the experiment, which will probably measure brain magnetic field details, and the precise check of its seriousness will not be simple.

Challenge 200, page 170: The maximum electric and magnetic field values are those that exert the maximum possible force $c^4/4G$ on an elementary charge e.

Challenge 202, page 171: See challenge 28.

Challenge 203, page 172: The electrons move slowly, but the speed of electrical signals is not given by the electron speed, but by the time at which the electrons move. Imagine long queue of cars (representing electrons) waiting in front of a red traffic light. All drivers look at the light. As soon as it turns green, everybody starts driving. Even though the driving speed might be only 10 m/s, the speed of traffic flow onset was that of light. It is this latter speed which is the speed of electrical signals.

Water pipes tell the same story. A long hose provides water almost in the same instant as the tap is opened, even if the water takes a long time to arrive from the tap to the end of the hose. The speed with which the water reacts is gives by the speed for pressure waves in water. Also for water hoses the signal speed, roughly given by the sound speed in water, is much higher than the speed of the water flow.

Challenge 204, page 172: One can measure smallest charges, showing that they are always multiples of the same unit. This method was used by Millikan. One can also measure current fluctuations, and show that they follow from shot noise, i.e., from the flow of discrete charges.

Challenge 207, page 173: Earth's potential would be $U = -q/(4\pi\varepsilon_o R) = 60$ MV, where the number of electrons in water must be taken into account.

Challenge 208, page 173: There is always a measurement error when measuring field values, even when measuring a 'vanishing' electromagnetic field. In addition, quantum theory leads to arbitrary small charge density values through the probability density due to wave functions.

Challenge 212, page 176: The issue is: is the 'universe' a concept? In the last volume of this ad-
venture we will show that it is not.

Challenge 214, page 181: When thinking, physical energy, momentum and angular momentum
are conserved, and thermodynamic entropy is not destroyed. Any experiment showing anything
different would point to unknown processes. However, there is no evidence for such processes.

Challenge 215, page 182: The best method cannot be much shorter than what is needed to de-
scribe 1 in 6000 million, or 33 bits. The Dutch and UK post code systems (including the letters
NL or UK) are not far from this value and thus can claim to be very efficient.

Challenge 216, page 182: For complex systems, when the unknowns are numerous, the advance
is thus simply given by the increase in answers. For the universe as a whole, the number of open
issues is quite low, as shown on page 238; here there has not been much advance in the last years.
But the advance is clearly measurable in this case as well.

Challenge 217, page 182: Is it possible to use the term 'complete' when describing nature? For a
clear-cut answer, see the last volume of our adventure.

Challenge 219, page 184: There are many baths in series: thermal baths in each light-sensitive
cell of the eyes, thermal baths inside the nerves towards the brain and thermal baths inside brain
cells.

Challenge 221, page 184: Yes.

Challenge 224, page 192: Physicists claim that the properties of objects, of space-time and of
interactions form the smallest list possible. However, this list is longer than the one found by
linguists! The reason is that physicists have found primitives that do not appear in everyday life.
In a sense, the aim of physicists is limited by list of unexplained questions of nature, given on
page 238.

Challenge 225, page 194: Neither has a defined content, clearly stated limits or a domain of
application.

Challenge 226, page 194: Impossible! That would not be a concept, as it has no content. The
solution to the issue must be and will be different.

Challenge 227, page 195: To neither. This paradox shows that such a 'set of all sets' does not
exist.

Challenge 228, page 196: The most famous is the class of all sets that do not contain themselves.
This is not a set, but a class.

Challenge 229, page 196: Dividing cakes is difficult. A simple method that solves many – but
not all – problems among N persons P1...PN is the following:

— P1 cuts the cake into N pieces.
— P2 to PN choose a piece.
— P1 keeps the last part.
— P2...PN assemble their parts back into one.
— Then P2...PN repeat the algorithm for one person less.

The problem is much more complex if the reassembly is not allowed. A *just* method (in finite
many steps) for 3 people, using nine steps, was published in 1944 by Steinhaus, and a *fully satis-
factory* method in the 1960s by John Conway. A fully satisfactory method for four persons was
found only in 1995; it has 20 steps.

Challenge 230, page 197: $(x, y) := \{x, \{x, y\}\}$.

Challenge 231, page 198: Hint: show that any countable list of reals misses at least one number.
This was proven for the first time by Cantor. His way was to write the list in decimal expansion

and then find a number that is surely not in the list. Second hint: his world-famous trick is called the diagonal argument.

Challenge 232, page 198: Hint: all reals are limits of series of rationals.

Challenge 234, page 199: Yes, but only provided division by zero is not allowed, and numbers are restricted to the rationals and reals.

Challenge 235, page 200: There are infinitely many of them. But the smallest is already quite large: 1016949152542372881355932203389830508474576271186440677966. If the number six is changed in the puzzle, one finds that the smallest solution for 1 is 1, for 4 is 102564, for 5 is 142857, for 8 is 1012658227848, for 2 is 105263157894736842, for 7 is 1014492753623188405797, for 3 is 1034482758620689655172413793, and for 9 is 10112359550561797752808988764044943820224719. The smallest solution for 6 is the largest of this list.

Challenge 236, page 200: $0 := \emptyset$, $1 := \{\emptyset\}$, $2 := \{\{\emptyset\}\}$ etc.

Challenge 237, page 204: Subtraction is easy. Addition is not commutative only for cases when infinite numbers are involved: $\omega + 2 \neq 2 + \omega$.

Challenge 238, page 204: Examples are $1 - \varepsilon$ or $1 - 4\varepsilon^2 - 3\varepsilon^3$.

Challenge 239, page 205: The answer is 57; the cited reference gives the details.

Challenge 240, page 207: $2^{2^{2^2}}$ and $4^{4^{4^4}}$.

Challenge 242, page 207: The child is minus 0.75 years old, or minus 9 months old; the father is thus very near the mother.

Challenge 243, page 207: This is not an easy question. The first non-trivial numbers are 7, 23, 47, 59, 167 and 179. See ROBERT MATTHEWS, Maximally periodic reciprocals, *Bulletin of the Institute of Mathematics and its Applications* **28**, pp. 147–148, 1992. Matthews shows that a number n for which $1/n$ generates the maximum of $n - 1$ decimal digits in the decimal expansion is a special sort of prime number that can be deduced from the so-called *Sophie Germain primes S*; one must have $n = 2S + 1$, where both S and $2S + 1$ must be prime and where S mod 20 must be 3, 9, or 11.

Thus the first numbers n are 7, 23, 47, 59, 167 and 179, corresponding to values for S of 3, 11, 23, 29, 83 and 89. In 1992, the largest known S that meets the criteria was

$$S = (39051 \cdot 2^{6002}) - 1 , \tag{111}$$

a 1812-digit long Sophie Germain prime number that is 3 mod 20. It was discovered by Wilfred Keller. This Sophie Germain prime leads to a prime n with a decimal expansion that is around 10^{1812} digits long before it starts repeating itself. Read your favourite book on number theory to find out more. Interestingly, the solution to this challenge is also connected to that of challenge 235. Can you find out more?

Challenge 244, page 207: Klein did not belong to either group. As a result, some of his nastier students concluded that he was not a mathematician at all.

Challenge 245, page 207: A barber cannot belong to either group; the definition of the barber is thus contradictory and has to be rejected.

Challenge 246, page 207: See the members.shaw.ca/hdhcubes/cube_basics.htm web page for more information on magic cubes.

Challenge 247, page 208: Such an expression is derived with the intermediate result $(1 - 2^2)^{-1}$. The handling of divergent series seems absurd, but mathematicians know how to give the expression a defined content. (See GODFREY H. HARDY, *Divergent Series*, Oxford University Press,

1949.) Physicists often use similar expressions without thinking about them, in quantum field theory.

Challenge 249, page 218: 'All Cretans lie' is *false*, since the opposite, namely 'some Cretans say the truth' is true in the case given. The trap is that the opposite of the original sentence is usually, but *falsely*, assumed to be 'all Cretans say the truth'.

Challenge 250, page 218: The statement cannot be false, due to the first half and the 'or' construction. Since it is true, the second half must be true and you are an angel.

Challenge 256, page 220: This can be debated; in any case it is definitely known that both statements are lies, as shown later on..

Challenge 257, page 220: If this false statement were true, swimmers or divers would also die, as their skin cannot breathe either.

Challenge 259, page 220: There is no way to settle this.

Challenge 262, page 221: The light bulb story seems to be correct. The bulb is very weak, so that the wire is not evaporating.

Challenge 263, page 221: The origin might be the number of people present in the last supper in the New Testament; or the forgotten 13th sign of the Zodiac. There is no truth in this superstition. In fact, every superstition is a lie. However, beware of people who are jealous of those who do not care about superstitions, and who get violent as a reaction.

Challenge 264, page 221: Without exception so far, all those who pretend to have been stigmatized have wounds in the *palms* of their hands. However, in crucifixion, the nails are driven through the *wrist*, because nails driven through the palms cannot carry the weight of a human body: the palms would tear open.

Challenge 265, page 221: The term 'multiverse' is both a belief and a lie. First of all, it is pure nonsense.

Challenge 267, page 221: In which frame of reference? How? Beware of anybody making that statement: he is a crook.

Challenge 272, page 227: Only induction allows us to make use of similarities and thus to define concepts.

Challenge 273, page 229: This depends on the definition (of the concept) of deity used.

Challenge 274, page 229: Yes, as we shall find out.

Challenge 275, page 230: Yes, as observation implies interaction.

Challenge 276, page 230: Lack of internal contradictions means that a concept is valid as a thinking tool; as we use our thoughts to describe nature, mathematical existence is a specialized version of physical existence, as thinking is itself a natural process. Indeed, mathematical concepts are also useful for the description of the working of computers and the like.

Another way to make the point is to stress that all mathematical concepts are built from sets and relations, or some suitable generalizations of them. These basic building blocks are taken from our physical environment. Sometimes the idea is expressed differently; many mathematicians have acknowledged that certain mathematical concepts, such as natural numbers, are taken directly from experience.

Challenge 277, page 230: Examples are Achilles, Odysseus, Mickey Mouse, the gods of polytheism and spirits.

Challenge 279, page 232: Torricelli made vacuum in a U-shaped glass tube, using mercury, the same liquid metal used in thermometers. Can you imagine how? A more difficult question: where did he get mercury from?

Challenge 280, page 233: Stating that something is infinite can be allowed, if the statement is falsifiable. An example is the statement 'There are infinitely many mosquitoes.'

Other statements are not falsifiable, such as 'The universe continue without limit behind the horizon.' Such a statement is a belief, not a fact.

Challenge 281, page 235: They are not sets either and thus not collections of points.

Challenge 282, page 235: There is still no possibility to interact with all matter and energy, as this includes oneself.

Challenge 283, page 241: No. There is only a generalization encompassing the two.

Challenge 284, page 242: An explanation of the universe is not possible, as the term explanation require the possibility to talk about systems outside the one under consideration. The universe is not part of a larger set.

Challenge 285, page 242: Both can in fact be seen as two sides of the same argument: there is no other choice;there is only one possibility. The rest of nature shows that it has to be that way, as everything depends on everything.

Challenge 289, page 259: The formula with $n - 1$ is a better fit. Why?

Challenge 292, page 264: The slowdown goes *quadratically* with time, because every new slowdown adds to the old one!

Challenge 293, page 264: No, only properties of parts of the universe are listed. The universe itself has no properties, as shown in the last volume..

Challenge 295, page 303: This could be solved with a trick similar to those used in the irrationality of each of the two terms of the sum, but nobody has found one.
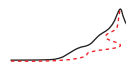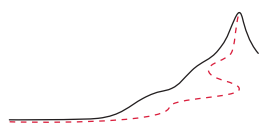
BIBLIOGRAPHY

> " [...] moi, qui trouve toujours tous les livres trop
> longs, et surtout les miens [...]
> Voltaire, *Lettre à M. Cideville.**

1 Julian Schwinger, L. L. DeRaad, K. A. Milton & W. Y. Tsai, *Classical Electrodynamics*, Perseus, 1998. An excellent text on the topic by one of its greatest masters.
See also the beautiful problem book by André Butoli & Jean-Marc Lévy-Leblond, *La physique en questions – électricité et magnétisme*, Vuibert, 1999. Cited on pages 15 and 69.

2 A pretty book about the history of magnetism and the excitement it generates is James D. Livingston, *Driving Force – the Natural Magic of Magnets*, Harvard University Press, 1996. Cited on page 17.

3 R. Edwards, Filling station fires spark cars' recall, *New Scientist*, pp. 4–5, 4 March 1995. Cited on page 17.

4 S. Desmet, F. Orban & F. Grandjean, On the Kelvin electrostatic generator, *European Journal of Physics* 10, pp. 118–122, 1989. You can also find construction plans for it in various places on the internet. Cited on page 19.

5 For an etching of Franklin's original ringing rod, see E. P. Krider, Benjamin Franklin and lightning rods, *Physics Today* 59, pp. 42–48, 2006. Cited on page 21.

6 W. Rueckner, An improved demonstration of charge conservation, *American Journal of Physics* 75, pp. 861–863, 2007. Cited on page 23.

7 For more details on various electromagnetic units, see the standard text by J. D. Jackson, *Classical Electrodynamics*, 3rd edition, Wiley, 1998. Cited on page 25.

8 See the old but beautiful papers by Richard C. Tolman & T. Dale Stewart, The electromotive force produced by the acceleration of metals, *Physical Review* 8, pp. 97–116, 1916, Richard C. Tolman & T. Dale Stewart, The mass of the electric carrier in copper, silver and aluminium, *Physical Review* 9, pp. 164–167, 1917, and the later but much more precise experiment by C. F. Kettering & G. G. Scott, Inertia of the carrier of electricity in copper and aluminum, *Physical Review* 66, pp. 257–267, 1944. (Obviously the American language dropped the 'i' from aluminium during that period.) The first of these papers is also a review of the preceding attempts, and explains the experiment in detail. The last paper shows what had to be taken into consideration to achieve sufficient precision. Cited on page 29.

9 This effect has first been measured by S. J. Barnett, A new electron-inertia effect and the determination of m/e for the free electron in copper, *Philosophical Magazine* 12, p. 349, 1931. Cited on page 29.

---

* '[...] me, who always finds all books too long, first of all my own [...]'.

10   See for example C. Schiller, A. A. Koomans,   van Rooy, C. Schönenberger & H. B. Elswijk, Decapitation of tungsten field emitter tips during sputter sharpening, *Surface Science Letters* **339**, pp. L925–L930, 1996. Cited on page 29.

11   L. I. Schiff & M. V. Barnhill, Gravitational-induced electric field near a metal, *Physical Review* **151**, pp. 1067–1071, 1966. F. C. Witteborn & W. M. Fairbank, Experimental comparison of the gravitational force on freely falling electrons and metallic electrons, *Physical Review Letters* **19**, pp. 1049–1052, 1967. Cited on page 30.

12   J. Lepak & M. Crescimanno, Speed of light measurement using ping, electronic preprint available at arxiv.org/abs/physics/0201053. Cited on page 30.

13   This story was printed on its fron page by the *Wall Street Journal* on 15 December 2006 under the title *Firms seek edge through speed as computer trading expands*. Cited on page 30.

14   See the website of the Particle Data Group, the world's reference, at pdg.web.cern.ch. See also H. Jeon & M. Longo, Search for magnetic monopoles trapped in matter, *Physical Review Letters* **75**, pp. 1443–1447, 1995. See also A. S. Goldhaber & W. P. Trower, Resource letter MM-1: magnetic monopoles, *American Journal of Physics* **58**, pp. 429–439, 1990. Cited on page 32.

15   Pierre de Maricourt, *Tractatus de magnete*, 1269. Cited on page 34.

16   R. Wiltschko & W. Wiltschko, *Magnetic Orientation in Animals*, Springer, Berlin, 1995. Cited on page 37.

17   The ratio of angular L to magnetic M moment is

$$\frac{L}{M} = \frac{2m}{e} \cdot \frac{1}{g} \; , \tag{112}$$

where e is the electron charge and m its mass. Both L and M are measurable. The first measurements were published with a g-value of 1, most probably because the authors expected the value. In later experiments, de Haas found other values. Measurements by other researchers gave values nearer to 2 than to 1, a fact that was only understood with the discovery of spin. The original publications are A. Einstein & W. J. de Haas, Proefondervinderlijk bewijs voor het bestaan der moleculaire stroomen van Ampère, *Konninklijke Akademie der Wetenschappen te Amsterdam, Verslagen* **23**, p. 1449, 1915, and A. Einstein & W. J. de Haas, Experimental proof of the existence of Ampère's molecular currents, *Konninklijke Akademie der Wetenschappen te Amsterdam, Proceedings* **18**, p. 696, 1916. Cited on page 39.

18   S. J. Barnett, Magnetization by rotation, *Physical Review* **6**, pp. 171–172, 1915, and S. J. Barnett, Magnetization by rotation, *Physical Review* **6**, pp. 239–270, 1915. Cited on page 40.

19   See J. D. Jackson, *Classical Electrodynamics*, 3rd edition, Wiley, 1998, or also R. F. Harrington, *Time Harmonic Electromagnetic Fields*, McGraw–Hill, 1961. Cited on pages 44 and 69.

20   The best book on the brain is Eric R. Kandel, James H. Schwartz & Thomas M. Jessell, *Principles of Neural Science*, fifth edition, McGraw-Hill, 2000. The suhep.phy.syr.edu/courses/modules/MM/Biology/biology2.html website gives an introduction into brain physiology. Cited on page 44.

21   N. Salingaros, Invariants of the electromagnetic field and electromagnetic waves, *American Journal of Physics* **53**, pp. 361–363, 1985. Cited on page 45.

22   A. C. DE LA TORRE, v ⩽ c in 1820?, *European Journal of Physics* **20**, pp. L23–L24, March 1999. Cited on page 46.

23   R. H. TYLER, S. MAUS & H. LÜHR, Magnetic signal due to ocean tidal flow identified in satellite observations, *Science* **299**, pp. 239–241, 2003. The films derived from the data can be found on the www.tu-bs.de/institute/geophysik/spp/publikationen.html website. Cited on page 53.

24   H. MONTGOMERY, Unipolar induction: a neglected topic in the teaching of electromagnetism, *European Journal of Physics* **20**, pp. 271–280, 1999. Cited on page 54.

25   On the geodynamo status, see the article by P. H. ROBERTS & G. A. GLATZMAIER, Geodynamo theory and simulations, *Reviews of Modern Physics* **72**, pp. 1081–1123, 2000. An older article is R. JEANLOZ & B. ROMANOWICZ, Geophysical dynamics at the center of the Earth, *Physics Today* pp. 22–27, August 1997. Cited on pages 54 and 152.

26   A. YAZDANI, D. M. EIGLER & N. D. LANG, Off-resonance conduction through atomic wires, *Science* **272**, pp. 1921–1924, 28 June 1996. For aluminium, gold, lead, niobium, as well as the influence of chemical properties, see ELKE SCHEER, The signature of chemical valence in the electric conduction through a single-atom contact, *Nature* **394**, pp. 154–157, 9 July 1998. Cited on page 57.

27   J. YANG, F. LU, L. W. KOSTIUK & D. Y. KWOK, Electrokinetic microchannel battery by means of electrokinetic and microfluidic phenomena, *Journal of Micromechanics and Microengineering* **13**, pp. 963–970, 2003. Cited on page 58.

28   See L. KOWALSKI, A myth about capacitors in series, *The Physics Teacher* **26**, pp. 286–287, 1988, and A. P. FRENCH, Are the textbook writers wrong about capacitors?, *The Physics Teacher* **31**, pp. 156–159, 1993. Cited on page 59.

29   G. G. P. VAN GORKUM, Introduction to Zeus displays, *Philips Journal of Research* **50**, pp. 269–280, 1996. See also N. LAMBERT, E. A. MONTIE, T. S. BALLER, G. G. P. VAN GORKUM, B. H. W. HENDRIKS, P. H. F. TROMPENAARS & S. T. DE ZWART, Transport and extraction in Zeus displays, *Philips Journal of Research* **50**, pp. 295–305, 1996. Cited on page 61.

30   A. L. HODGKIN & A. F. HUXLEY, A quantitative description of membrane current and its application to conduction and excitation in nerve, *Journal of Physiology* **117**, pp. 500–544, 1952. This famous paper of theoretical biology earned the authors the Nobel Prize in Medicine in 1963. Cited on page 61.

31   See the excellent overview article by T. HEIMBURG, Die Physik von Nerven, *Physik Journal* **8**, pp. 33–39, 2009. See also S. S. L. ANDERSEN, A. D. JACKSON & T. HEIMBURG, Towards a thermodynamic theory of nerve pule propagation, *Progress in Neurobiology* **88**, pp. 104–113, 2009, the website membranes.nbi.dk, and the text THOMAS HEIMBURG, *Thermal Biophysics of Membranes*, Wiley-VCH, 2007. Cited on page 62.

32   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 63 and 169.

33   A similar summary is the basis of FRIEDRICH W. HEHL & YURI N. OBUKOV, *Foundations of Classical Electrodynamics – Charge, Flux and Metric*, Birkhäuser 2003. Cited on page 63.

34   OLEG D. JEFIMENKO, A relativistic paradox seemingly violating conservation of momentum law in electromagnetic systems, *European Journal of Physics* **20**, pp. 39–44, 1999. Of course, the missing momentum goes into the electromagnetic field. Given that the electromagnetic momentum is given by the vector potential, are you able to check whether everything comes out right? Cited on page 68.

35 H. Van Dam & E. P. Wigner, Classical relativistic mechanics of interacting point particles, *Physical Review* **136B**, pp. 1576–1582, 1965. Cited on page 68.

36 Mark D. Semon & John R. Taylor, Thoughts on the magnetic vector potential, *American Journal of Physics* **64**, pp. 1361–1369, 1996. Cited on pages 69 and 71.

37 Jean Sivardière, Simple derivation of magnetic vector potentials, *European Journal of Physics* **14**, pp. 251–254, 1993. Cited on page 70.

38 T. T. Wu & C. N. Yang, 1975, Concept of nonintegrable phase factors and global formulation of gauge fields, *Physical Review D* **12**, pp. 3845–3857, Cited on page 73.

39 An electrodynamics text completely written with (mathematical) forms is Kurt Meetz & Walter L. Engl, Elektromagnetische Felder – mathematische und physikalische Grundlagen, Springer, 1980. Cited on page 74.

40 J. Travis, Twirl those organs into place – getting to the heart of how a heart knows left from right, *Science News* **156**, 21 August, 1999. A good book on asymmetries in nature is H. Brunner, *Rechts oder links*, Wiley–Vch, 1999. Cited on page 77.

41 See for example the discussion by M. C. Corballis & I. L. Beale, On telling left from right, *Scientific American* **224**, pp. 96–104, March 1971. Cited on page 77.

42 In 1977, Claus Montonen and David Olive showed that quantum theory allows duality transformations even *with* the inclusion of matter, if specific types of magnetic monopoles, the so-called *dyons*, exist. The fundamental paper is D. Olive & C. Montonen, Magnetic monopoles as gauge particles, *Physics Letters* **72B**, pp. 117–120, 1977. Many other papers built on this one; however, no experimental support for the approach has ever appeared. Cited on page 78.

43 Wolfgang Rindler, *Essential Relativity – Special, General, and Cosmological*, revised 2nd edition, Springer Verlag, 1977, page 247. There is also the beautiful paper by M. Le Bellac & J.-M. Lévy-Leblond, Galilean electrodynamics, *Nuovo Cimento B* **14**, p. 217, 1973, that explains the possibilities but also the problems appearing when trying to define the theory non-relativistically. Cited on page 79.

44 L.-C. Tu, J. Luo & G. T. Gilles, The mass of the photon, *Reports on Progress of Physics* **68**, pp. 77–130, 2005. Cited on page 79.

45 The system for typing by thought alone is described in many papers, such as B. Blankertz, F. Losch, M. Krauledat, G. Dornhege, G. Curio & K.-R. Müller, The Berlin Brain-Computer Interface: accurate performance from first session in BCI-naïve subjects, *IEEE Transactions on biomedial engineering* **55**, pp. 2452–2462, 2008. See the website www.bbci.de for more information. Cited on page 80.

46 D. Singleton, Electromagnetic angular momentum and quantum mechanics, *American Journal of Physics* **66**, pp. 697–701, 1998, Cited on page 81.

47 For a captivating account on the history of the ideas on light, see David Park, *The Fire Within the Eye: a Historical Essay on the Nature and Meaning of Light*, Princeton University Press, 1997. For an example of the complex history of optics, see the famous text by Alhazen or Ibn al-Haytham, *Book of Optics* 1021. However, no Arabic website allows reading the text, and the Arabic Wikipedia articles on the topic are much shorter than the French or English ones. Indeed, like most ancient Middle-East thinkers, Alhazen is better known in Europe than in his home region. A Latin translation of the Book of Optics can be read at the imgbase-scd-ulp.u-strasbg.fr/displayimage.php?album=44&pos=0 website of the Université de Strasbourg. Cited on page 83.

48 See the text by Raymond L. Lee & Alistair B. Fraser, *The Rainbow Bridge: Rainbows in Art, Myth, and Science*, Pennsylvania State University Press, 2000. A chapter can be

found at the www.nadn.navy.mil/Oceanography/RainbowBridge/Chapter_8.html website. Cited on page 88.

49 The beautiful slit experiment was published by E. A. MONTIE, E. C. COSMAN, G. W. 'T HOOFT, VAN DER MARK & C. W. J. BEENAKKER, Observation of the optical analogue of quantized conductance of a point contact, *Nature* **350**, pp. 594–595, 18 April 1991, and in the longer version E. A. MONTIE, E. C. COSMAN, G. W. 'T HOOFT, M. B. VAN DER MARK & C. W. J. BEENAKKER, Observation of the optical analogue of the quantised conductance of a point contact, *Physica B* **175**, pp. 149–152, 1991. The result was also publicized in numerous other scientific magazines. Cited on page 88.

50 A recent measurement of the frequency of light is presented in TH. UDEM, A. HUBER, B. GROSS, J. REICHERT, M. PREVEDELLI, M. WEITZ & T. W. HAUSCH, Phase-coherent measurement of the hydrogen 1S–2S transition frequency with an optical frequency interval divider chain, *Physical Review Letters* **79**, pp. 2646–2649, 6 October 1997. Another is C. SCHWOB, L. JOZEFOWSKI, DE BEAUVOIR, L. HILICO, F. NEZ, L. JULIEN, F. BIRABEN, O. ACEF & A. CLAIRON, Optical frequency measurement of the 2S-12D transitions in hydrogen and deuterium: Rydberg constant and Lamb shift determinations, *Physical Review Letters* **82**, pp. 4960–4963, 21 June 1999. Cited on page 89.

51 The discoverors of two such methods were awarded the 2005 Nobel Prize for physics. Cited on page 89.

52 The best introduction to mirages are the web pages by Andrew Young at mintaka.sdsu.edu/GF/mintaka/mirintro.html. He explains the many types that exist: inferior mirages, superior mirages, mock mirages, Wegener-type mirages, etc., and also gives many references, clearly showing which ones give correct and which one give incorrect explanations. There is no modern review article on the topic yet. For a curious aspect, see G. HORVÁTH, J. GÁL & R. WEHNER, Why are water-seeking insects not attracted by mirages? The polarization pattern of mirages, *Naturwissenschaften* **83**, pp. 300–303, 1997. Cited on page 90.

53 W. K. HAIDINGER, Über das direkte Erkennen des polarisierten Lichts, *Poggendorf's Annalen* **63**, pp. 29–39, 1844, W. K. HAIDINGER, Beobachtung des Lichtpolarisationsbüschels in geradlinig polarisiertem Lichte, *Poggendorf's Annalen* **68**, pp. 73–87, 1846, W. K. HAIDINGER, Dauer des Eindrucks der Polarisationsbüschel auf der Netzhaut, *Poggendorf's Annalen* **93**, pp. 318–320, 1854. Cited on page 90.

54 See the chapter on polarization brushes in MARCEL G. J. MINNAERT, *Light and Colour in the Outdoors*, Springer, 1993, or the original book series, MARCEL G. J. MINNAERT, *De natuurkunde van 't vrije veld*, Thieme & Cie, 1937. For more details, see G. P. MISSION, Form and behaviour of Haidinger's brushes, *Ophthalmology and Physiological Optics* **137**, pp. 392–396, 1993, or J. GREBE-ELLIS, Zum Haidinger-Büschel, 2002, at www.physik.hu-berlin.de/ger/gruppen/didaktik/forschung/optik/grebe_ellis_dpg2002.pdf. On the birefringence of the eye, see L. BOUR, Een eigenaardige speling der natuur, *Nederlands tijdschrift voor natuurkunde* **67**, pp. 362–364, December 2001. In particular, a photograph of the eye using linear polarized illumination and taken through an analyser shows a black *cross* inside the pupil. Cited on page 90.

55 M. BURRESI, D. VAN OSTEN, T. KAMPFRATH, H. SCHOENMAKER, R. HEIDEMAN, A. LEINSE & L. KUIPERS, Probing the magnetic field of light at optical frequencies, *Science Express* October 2009. Cited on page 93.

56 K. L. KELLY, Color designations for colored lights, *Journal of the Optical Society of America* **33**, pp. 627–632, 1943. Cited on page 95.

57 EDWARD M. PURCELL, *Electricity and Magnetism – Berkeley Physics Course Volume 2*, McGraw–Hill, 1984. Cited on page 97.

58 This was the book series in twenty volumes by A ARON B ERNSTEIN, *Naturwissenschaftliche Volksbücher*, Duncker, 1873-1874. The young Einstein read them, between 1892 and 1894, with 'breathless attention', as he wrote later on. Cited on page 99.

59 The first correct explanation of the light mill was given by O SBORNE R EYNOLDS, On certain dimensional properties of matter in the gaseous state, *Royal Society Philosophical Transactions* Part 2, 1879. The best discussion is the one given on the web by P HIL G IBBS, in the frequently asked question list of the usenet news group sci.physics; it is available at the www.desy.de/user/projects/Physics/light-mill.html website. Cited on page 102.

60 P. L EBEDEW, Untersuchungen über die Druckkräfte des Lichtes, *Annalen der Physik* **6**, pp. 307–458, 1901. Lebedew confirmed Kepler's result that light pressure is the basis for the change of direction of the tails of comets when they circle around the Sun. Cited on page 102.

61 P. G ALAJDA & P. O RMOS, *Applied Physics Letters* **78**, p. 249, 2001. Cited on page 102.

62 A short overview is given by M ILES P ADGETT & L ES A LLEN, Optical tweezers and spanners, *Physics World* pp. 35–38, September 1997. The original papers by Ashkin's group are A. A SHKIN, J. M. D ZIEDZIC, J. E. B JORKHOLM & S. C HU, Observation of a gradient force optical trap for dielectric particles, *Optics Letters* **11**, p. 288, 1986, and A. A SKIN, J. M. D ZIEDZIC & T. Y AMANE, Optical trapping and manipulation of single cells using infrared laser beams, *Nature* **330**, p. 769, 1987. A pedagogical explanation on optical spanners, together with a way to build one, can be found in D. N. M OOTHOO, J. A RLT, R. S. C ONROY, F. A KERBOOM, A. V OIT & K. D HOLAKIA, Beth's experiment using optical tweezers, *American Journal of Physics* **69**, pp. 271–276, 2001, and in S. P. S MITH, S. R. B HALOTRA, A. L. B RODY, B. L. B ROWN, E. K. B OYDA & M. P RENTISS, Inexpensive optical tweezers for undergraduate laboratories, *American Journal of Physics* **67**, pp. 26–35, 1999. Cited on page 102.

63 R. A. B ETH, Mechanical detection and measurement of the angular momentum of light, *Physical Review* **50**, p. 115, 1936. For modern measurements, see N. B. S IMPSON, K. D HOLAKIA, L. A LLEN & M. J. P ADGETT, Mechanical equivalence of spin and orbital angular momentum of light:an optical spanner, *Optics Letters* **22**, pp. 52–54, 1997, and M. E. J. F RIESE, T. A. N IEMINEN, N. R. H ECKENBERG & H. R UBINSZTEIN-D UNLOP, Optical torque controlled by elliptical polarization, *Optics Letters* **23**, pp. 1–3, 1998. See also J. H. P OYNTING, The wave motion of a revolving shaft, and a suggestion as to the angular momentum in a beam of circularly polarised light, *Proceedings of the Royal Society London A* **82**, pp. 560–567, 1908. Cited on page 102.

64 A. V ALENZUELA, G. H AERENDEL, H. F ÖPPL, F. M ELZNER, H. N EUSS, E. R IEGER, J. S TÖCKER, O. B AUER, H. H ÖFNER & J. L OIDL, The AMPTE artificial comet experiments, *Nature* **320**, pp. 700–703, 1986. Cited on page 103.

65 See the Latin text by D IETRICH VON F REIBERG, *De iride et radialibus impressionibus*, c. 1315. Cited on page 105.

66 J. W ALKER, Multiple rainbows from single drops of water and other liquids, *American Journal of Physics* **44**, pp. 421–433, 1976, and his How to create and observe a dozen rainbows in a single drop of water, *Scientific American* **237**, pp. 138–144, 1977. See also K. S ASSEN, Angular scattering and rainbow formation in pendant drops, *Journal of the Optical Society of America* **69**, pp. 1083–1089, 1979. Cited on page 105.

67 The resulting colouring of the Sun's rim is shown clearly on Andrew Young's web page mintaka.sdsu.edu/GF/explain/simulations/std/rims.html. His website mintaka.sdsu.edu/GF offers the best explanation of the green flash, including the various types that exist (ex-

plained at mintaka.sdsu.edu/GF/papers/Zenit/glance.html), how to observe it, and the numerous physical effects involved. Detailed simulations and extensive material is available. See also his paper A. T. Young, Sunset science – III. Visual adaptation and green flashes, *Journal of the Optical Society of America A* **17**, pp. 2129–2139, 2000. Cited on pages 105 and 106.

68 There are also other ways to see the green ray, for longer times, namely when a mirage appears at sunset. An explanation with colour photograph is contained in M. Vollmer, Gespiegelt in besonderen Düften …– Oasen, Seeungeheuer und weitere Spielereien der Fata Morgana, *Physikalische Blätter* **54**, pp. 903–909, 1998. Cited on page 105.

69 See the wonderful website by Les Cowley on atmospheric optics, www.atoptics.co.uk. Or the book David K. Lynch & William Livingston, *Color and Light in Nature*, second edition, Cambridge University Press, 2001. They updated and expanded the fascination for colours in nature – such as, for example, the halos around the Moon and the Sun, or the colour of shadows – that was started by the beautiful and classic book by Marcel Minnaert mentioned on page 86. Cited on page 106.

70 This famous discovery is by Brent Berlin & Paul Kay, *Basic Color Terms: Their Universality and Evolution*, University of California Press, 1969. Their ongoing *world colour survey* is eagerly awaited. Of course there are also ongoing studies to find possible exceptions; the basic structure is solid, as shown in the conference proceedings C. L. Hardin & Luisa Maffi, *Colour Categories in Thought and Language*, Cambridge University Press, 1997. Cited on page 107.

71 For a thorough discussion of the various velocities connected to wave trains, see the classic text by Louis Brillouin, *Wave Propagation and Group Velocity*, Academic Press, New York, 1960. It expands in detail the theme discussed by Arnold Sommerfeld, Über die Fortpflanzung des Lichtes in dispergierenden Medien, *Annalen der Physik,* 4th series, **44**, pp. 177–202, 1914. See also Arnold Sommerfeld, *Optik*, Dietrichssche Verlagsbuchandlung, Wiesbaden 1950, section 22. An English translation Arnold Sommerfeld, *Lectures on Theoretical Physics: Optics*, 1954, is also available. Cited on pages 108, 110, and 111.

72 Changing the group velocity in fibres is now even possible on demand, as shown by M. González-Herráez, K.-Y. Song & L. Thévenaz, Optically controlled slow and fast light in optical fibers using stimulated Brillouin scattering, *Applied Physics Letters* **87**, p. 081113, 2005. They demonstrate group velocities from $0.24c$ to plus infinity and beyond, to negative values.

Another experiment was carried out by S. Chu & S. Wong, Linear pulse propagation in an absorbing medium, *Physical Review Letters* **48**, pp. 738–741, 1982. See also S. Chu & D. Styer, Answer to question #52. Group velocity and energy propagation, *American Journal of Physics* **66**, pp. 659–661, 1998. Another example was described in 1993 by the group of Raymond Chiao for the case of certain nonlinear materials in R. Chiao, P. G. Kwait & A. M. Steinberg, Faster than light?, *Scientific American* **269**, p. 52, August 1993, and R. Y. Chiao, A. E. Kozhekin & G. Kurizki, Tachyonlike excitations in inverted two-level media, *Physical Review Letters* **77**, pp. 1254–1257, 1996. On still another experimental set-up using anomalous dispersion in caesium gas, see L. J. Wang, A. Kuzmich & A. Dogarin, Gain-assisted superluminal light propagation, *Nature* **406**, pp. 277–279, 20 July 2000. Cited on page 110.

73 Y. P. Terletskii, *Paradoxes in the Theory of Relativity*, Plenum Press, 1968. Cited on page 111.

74 See the excellent explanation by Kirk T. McDonald, Negative group velocity, *American Journal of Physics* **69**, pp. 607–614, 2001. Cited on page 111.

75 G. Nimtz, A. Enders & H. Spieker, *Journal de Physique I (Paris)* **4**, p. 565, 1994. Unfortunately, Nimtz himself seems to believe that he transported energy or signals faster than light; he is aided by the often badly prepared critics of his quite sophisticated experiments. See A. Enders & G. Nimtz, *Physikalische Blätter* **49**, p. 1119, Dezember 1993, and the weak replies in *Physikalische Blätter* **50**, p. 313, April 1994. See also A. M. Steinberg, *Journal de Physique I (Paris)* **4**, p. 1813, 1994, A. M. Steinberg, P. G. Kwiat & R. Y. Chiao, *Physical Review Letters* **71**, pp. 708–711, 1993, and A. Ranfagni, P. Fabeni, G. P. Pazzi & D. Mugnai, *Physical Review E* **48**, p. 1453, 1993. Cited on page 111.

76 A summary of all evidence about the motion of the aether is given by R. S. Shankland, S. W. McCuskey, F. C. Leone & G. Kuerti, New analysis of the interferometer observations of Dayton C. Miller, *Review of Modern Physics* **27**, pp. 167–178, 1955. An older text is H. Witte, *Annalen der Physik* **26**, p. 235, 1908. Cited on page 112.

77 The history of the concept of vacuum can be found in the book by E. Grant, *Much Ado About Nothing*, Cambridge University Press, 1981, and in the extensive reference text by Edmund T. Whittaker, *A History of the Theories of Aether and Electricity*, Volume 1: *The Classical Theories*, Volume 2: *The Modern Theories*, Tomash Publishers, American Institute of Physics 1951, 1987.

The various *aether models* – gears, tubes, vortices – proposed in the nineteenth century were dropped for various reasons. Since many models used to explain electric and magnetic fields as motion of some entities, it was concluded that the speed of light would depend on electric or magnetic fields. One type of field was usually described by linear motion of the entities, the other by rotary or twisting motion; both assignments are possible. As a consequence, aether must be a somewhat strange fluid that flows perfectly, but that resists rotation of volume elements, as McCullogh deduced in 1839. However, experiments show that the speed of light in vacuum does not depend on electromagnetic field intensity. Vortices were dropped because real world vortices were found out to be unstable. All models received their final blow when they failed to meet the requirements of special relativity. Cited on page 113.

78 See its www.cie.co.at/cie website. Cited on page 118.

79 P. D. Jones, M. New, D. E. Parker, S. Martin & I. G. Rigor, Surface air temperature and its changes over the past 150 years, *Reviews of Geophysics* **37**, pp. 173–199, May 1999. Cited on page 118.

80 Pictures of objects in a red hot oven and at room temperature are also shown in C. H. Bennett, Demons, engines and the second law, *Scientific American* **255**, pp. 108–117, November 1987. Cited on page 119.

81 If you want to read more on the topic, have a look at the classic text by Warren J. Smith, *Modern Optical Engineering : the Design of Optical Systems*, 3rd edition, McGraw-Hill, 2000. The main historic reference is R. Clausius, Über die Concentration von Wärme und Lichtstrahlen und die Gränzen ihrer Wirkung, *Poggendorff's Annalen der Physik* **121**, pp. 1–44, 1864. Cited on pages 120 and 144.

82 Measured values and ranges for physical quantities are collected in Horst Völz & Peter Ackermann, *Die Welt in Zahlen*, Spektrum Akademischer Verlag, 1996. Cited on page 121.

83 See e.g. K. Codling & L. J. Frasinski, Coulomb explosion of simple molecules in intense laser fields, *Contemporary Physics* **35**, pp. 243–255, 1994. Cited on page 121.

84 The standard reference on the propagation of light is Max Born & Emil Wolf, *Principles of Optics – Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Pergamon Press, 6th edition, 1998. Cited on page 123.

85 E. D. Palik, *Handbook of optical constants of solids*, Academic Publishing, 1998. Cited on page 125.

86 More mirage photographs, even mirage films, can be found on www.polarimage.fi/mirages/mirages.htm and virtual.finland.fi/netcomm/news/showarticle.asp?intNWSAID=25722. Cited on page 126.

87 E. J. J. Kirchner, De uitvinding van het telescoop in 1608: gewoon twee lenzen, *neder-lands tijdschrift voor natuurkunde* 74, pp. 356–361, 2008. Cited on page 126.

88 A fascinating overview about what people have achieved in this domain up to now is given by Peter Manly, *Unusual Telescopes*, Cambridge University Press, 1991, and Henry C. King, *The History of the Telescope*, Dover, 2003. Cited on page 127.

89 An introduction to the topic of the 22° halo, the 46° halo, Sun dogs, and the many other arcs and bows that can be seen around the Sun, see the beautifully illustrated paper by R. Greenler, Lichterscheinungen, Eiskristalle und Himmelsarchäologie, *Physikalische Blätter* 54, pp. 133–139, 1998, or the book Robert Greenler, *Rainbows, Halos, and Glo-ries*, Cambridge University Press, 1980. Cited on page 128.

90 K. Franz & al., Müller cells are living optical fibers in the vertebrate retina, *Proceedings of the National Academy of Sciences* 104, pp. 8287–8292, 2007. Cited on page 130.

91 A complete list of data and arguments showing that the hair of polar bears have no fibre function is found on the pages it.stlawu.edu/~koon/mar-ref.html and it.stlawu.edu/~koon/polar.html. Cited on page 130.

92 The prediction of negative refraction is due to V. G. Veselago, The electrodynamics of substances with simultaneously negative values of ε and μ, *Soviet Physics Uspekhi* 10, p. 509, 1968. (The original paper in Russian is from 1967.) The explanation with different refraction directions was published by P. M. Valanju, R. M. Walser & A. P. Valanju, Wave re-fraction in negative-index media: always positive and very inhomogeneous, *Physical Review Letters* 88, p. 187401, 8 May 2002. Also Fermat's principle is corrected, as explained in V. G. Veselago, About the wording of Fermat's principle for light propagation in media with negative refraction index, arxiv.org/abs/cond-mat/0203451. Cited on page 130.

93 The first example of material system with a negative refraction index were presented by David Smith and his team. R. A. Schelby, D. R. Smith & S. Schultz, Experimental verification of a negative index of refraction, *Science* 292, p. 77-79, 2001. More recent ex-amples are A. A. Houck, J. B. Brock & I. L. Chuang, Experimental observations of a left-handed material that obeys Snell's law, *Physical Review Letters* 90, p. 137401, 2003, C. G. Parazzoli, R. B. Greegor, K. Li, B. E. C. Koltenbah & M. Tanielian, Experimental verification and simulation of negative index of refraction using Snell's law, *Physical Review Letters* 90, p. 107401, 2003. S. Foteinopoulou, E. N. Economou & C. M. Soukoulis, Refraction in media with a negative refractive index, *Physical Review Letters* 90, p. 107402, 2003. (Many call the 'law' of refraction 'Snell's law', even though many others found it before him, and even though his family name was 'Snel'.) Cited on page 131.

94 S. A. Ramakrishna, Physics of negative refractive index materials, *Reorts on Progress of Physics* 68, pp. 449–521, 2005. Cited on page 131.

95 J. Pendry, Negegative refraction makes a perfect lens, *Physical Review Letters* 85, p. 3966, 2000. See also J. B. Pendry, D. Schurig & D. R. Smith, Controlling electromagnetic fields, *Science* 312, pp. 1780–1782, 2006, and D. Schurig, J. J. Mock, B. J. Justice, S. A. Cummer, J. B. Pendry, A. F. Starr & D. R. Smith, Metamaterial electromag-netic cloak at microwave frequencies, *Science* 314, pp. 977–980, 2006. Cited on page 131.

96 On metamaterials, see A. Lai, C. Caloz & T. Itoh, Composite rightleft-handed trans-

mission metamaterials, *IEEE Microwave Magazine* **5**, pp. 34–50, September 2004. Cited on page 131.

97   M. ZEDLER & P. RUSSER, Investigation on the Dispersion Relation of a 3D LC-based Metamaterial with an Omnidirectional Left-Handed Frequency Band, *2006 International Microwave Symposium Digest, San Francisco* pp. 1477–1479, 2006. M. ZEDLER, C. CALOZ & P. RUSSER, A 3D Isotropic left-handed metamaterial based on the rotated transmission line matrix (TLM) scheme, *IEEE Transactions on Microwave Theory and Techniques* **55**, pp. 2930–2941, 2007. Cited on page 131.

98   Read Grimaldi's text online at fermi.imss.fi.it/rd/bdv?/bdviewer/bid=000000300682. Cited on page 132.

99   JAMES E. FALLER & JOSEPH WAMPLER, The lunar laser reflector, *Scientific American* pp. 38–49, March 1970. Cited on page 132.

100   Neil Armstrong of Apollo 11, Jim Lovell of Apollo 8 and Apollo 13, and Jim Irwin of Apollo 15 extensively searched for it and then made negative statements, as told in *Science News* p. 423, 24 & 31 December 1994. From the space shuttle however, which circles only a few hundred kilometres above the Earth, the wall *can* be seen when the Sun is low enough such that the wall appears wider through its own shadow, as explained in *Science News* **149**, p. 301, 1996. Cited on page 132.

101   S. W. HELL, Strategy for far-field optical imaging and writing without diffraction limit, *Physics Letters A* **326**, pp. 140–145, 2004, see also V. WESTPHAL & S. W. HELL, Nanoscale resolution in the focal plane of an optical microscope, *Physical Review Letters* **94**, p. 143903, 2005, and V. WESTPHAL, J. SEEGER, T. SALDITT & S. W. HELL, Stimulated emission depletion microscopy on lithographic microstructures, *Journal of Physics B* **38**, pp. S695–S705, 2005. Cited on page 134.

102   M. SHIH, M. SEGEV & G. SALAMO, *Physical Review Letters* **78**, pp. 2551–2554, 1997. See also the more readable paper by MORDECHAI SEGEV & GEORGE STEGEMAN, Self-trapping of optical beams: spatial solitons, *Physics Today* **51**, pp. 42–48, August 1998. Cited on page 134.

103   W. H. EHRENSTEIN & B. LINGELBACH, Das Hermann–Gitter, *Physik in unserer Zeit* **6**, pp. 263–268, 2002. The journal also shows a colour variation of these lattices. Cited on page 137.

104   To enjoy many other flowers under UV illumination, go to the extensive collection at www.naturfotograf.com/index2. Cited on page 140.

105   The eye sensitivity myth is debunked in detail by B. H. SOFFER & D. K. LYNCH, Some paradoxes, errors, and resolutions concerning the spectral optimization of human vision, *American Journal of Physics* **67**, pp. 946–953, 1999. Cited on page 140.

106   DAVID R. WILLIAMS, Supernormal Vision, *Science News* **152**, pp. 312–313, 15 November 1997. See also www.cvs.rochester.edu/people/d_williams/d_williams.html as well as the photographs at www.cvs.rochester.edu/people/~aroorda/ao_research.html of the interior of living human eyes. Their last publication is A. ROORDA, A. METHA, P. LENNIE & D. R. WILLIAMS, Packing arrangement of the three cone classes in the primate retina, *Vision Research* **41**, pp. 1291–1306, 2001. Cited on page 140.

107   This happened to Giovanni Bellini (*c.* 1430–1516) the great Venetian Renaissance painter, who even put this experience into writing, thus producing one of the greatest 'gaffes' ever. If you take a photograph of the effect with a remotely controlled camera, you can prove that your camera is holy as well. Cited on page 142.

108 S. R. Wɪʟᴋ, How retroreflectors really work, *Optics & Photonics News,* pp. 6–7, December 1993. Cited on page 142.

109 This problem was suggested by Vladimir Surdin. Cited on page 142.

110 Such a claim was implicitly made by D. Mᴜɢɴᴀɪ, A. Rᴀɴғᴀɢɴɪ & R. Rᴜɢɢɪᴇʀɪ, Observation of superluminal behaviors in wave propagation, *Physical Review Letters* **84**, p. 4830, 2000. An excellent explanation and rebuttal was given by W. A. Rᴏᴅʀɪɢᴜᴇs, D. S. Tʜᴏʙᴇʀ & A. L. Xᴀᴠɪᴇʀ, Causal explanation for observed superluminal behavior of microwave propagation in free space, arxiv.org/abs/physics/0012032. Cited on page 142.

111 If you want to see more on how the world looks for the different types of colour blind, have a look at the webexhibits.org/causesofcolor/2.html or the www.vischeck.com/examples web pages. Cited on page 142.

112 Most of the world's experts in lightning are Russian. Two good books are Vʟᴀᴅɪᴍɪʀ A. Rᴀᴋᴏᴠ & Mᴀʀᴛɪɴ A. Uᴍᴀɴ, *Lightning: Physics and Effects*, Cambridge University Press, 2003, and Eᴅᴜᴀʀᴅ M. Bᴀᴢᴇʟʏᴏɴ & Yᴜʀɪ P. Rᴀɪᴢᴇʀ, *Lightning Physics and Lightning Protection*, Institute of Physics Publishing, 2000. Cited on page 147.

113 About the life-long passion that drove Luke Howard, see the book by Rɪᴄʜᴀʀᴅ Hᴀᴍʙʟʏɴ, *The Invention of Clouds*, Macmillan 2001. Cited on page 147.

114 J. Lᴀᴛʜᴀᴍ, The electrification of thunderstorms, *Quartely Journal of the Royal Meteorological Society* **107**, pp. 277–289, 1981. For a more recent and wider review, see Eᴀʀʟᴇ R. Wɪʟʟɪᴀᴍs, The tripole structure of thunderstorms, *Journal of Geophysical Research* **94**, pp. 13151–13167, 1989. See also the book by the Nᴀᴛɪᴏɴᴀʟ Rᴇsᴇᴀʀᴄʜ Cᴏᴜɴᴄɪʟ Sᴛᴀғғ, *The Earth's Electrical Environment*, Studies in Geophysics, National Academy Press, 1986. Cited on page 147.

115 See, for example, the overview at enviromom.us/lightning/lightningformation.html. Cited on page 147.

116 A. V. Gᴜʀᴇᴠɪᴄʜ & K. P. Zʏʙɪɴ, Runaway breakdown and the mysteries of lightning, *Physics Today* **58**, pp. 37–43, May 2005. Cited on page 147.

117 To learn more about atmospheric currents, you may want to have a look at the popularizing review of the US work by E. A. Bᴇʀɪɴɢ, A. A. Fᴇᴡ & J. R. Bᴇɴʙʀᴏᴏᴋ, The global electric circuit, *Physics Today* **51**, pp. 24–30, October 1998, or the more technical overview by E. Bᴇʀɪɴɢ, *Reviews of Geophysics* (supplement) **33**, p. 845, 1995. Cited on page 150.

118 The use of Schumann resonances in the Earth–ionosphere capacitor for this research field is explained in K. Sᴄʜʟᴇɢᴇʟ & M. Fᴜ̈ʟʟᴇʀᴋʀᴜɢ, Weltweite Ortung von Blitzen, *Physik in unserer Zeit* **33**, pp. 257–261, 2002. Cited on page 150.

119 J. R. Dᴡʏᴇʀ, M. A. Uᴍᴀɴ, H. K. Rᴀssᴏᴜʟ, M. Aʟ-Dᴀʏᴇʜ, E. L. Cᴀʀᴀᴡᴀʏ, J. Jᴇʀᴀᴜʟᴅ, V. A. Rᴀᴋᴏᴠ, D. M. Jᴏʀᴅᴀɴ, K. J. Rᴀᴍʙᴏ, V. Cᴏʀʙɪɴ & B. Wʀɪɢʜᴛ, Energetic radiation produced by rocket-triggered lightning, *Science* **299**, pp. 694–697, 2003. Cited on page 150.

120 J. R. Dᴡʏᴇʀ, A fundamental limit on electric fields in air, *Geophysical Research Letters* **30**, p. 2055, 2003. Cited on page 150.

121 B. M. Sᴍɪʀɴᴏᴠ, Physics of ball lightning, *Physics Reports* **224**, pp. 151–236, 1993. See also D. Fɪɴᴋᴇʟsᴛᴇɪɴ & J. Rᴜʙɪɴsᴛᴇɪɴ, Ball lightning, *Physical Review* **135**, pp. 390–396, 1964. For more folklore on the topic, just search the world wide web. Cited on page 151.

122 G. D. Sʜᴀʙᴀɴᴏᴠ, The optical properties of long-lived luminous formations, *Technical Physics Letters* **28**, pp. 164–166, 2002, A. I. Eɢᴏʀᴏᴠ & S. I. Sᴛᴇᴘᴀɴᴏᴠ, Long-lived plasmoids produced in humid air as analogues of ball lightning, *Technical Physics* **47**, pp. 1584–

1586, 2002, A. E. Egorov, S. I. Stepanov & G. D. Shabanov, *Physics Uspekhi* Laboratory demonstration of ball lightning, 47, pp. 99–101, 2004, and G. D. Shabanov & B. Yu. Sokolovskii, Macroscopic separation of charges in a pulsed electric discharge, *Plasma Physics Reports* **31**, pp. 512–518, 2005. (All these are English translations of earlier Russian papers.) See the websites biod.pnpi.spb.ru/pages_ru/Stepanov/index.html stealthtank.narod.ru, balllightning.narod.ru/hvewd.html and www.ipp.mpg.de/ippcms/eng/presse/pi/05_06_pi.html, for more details and more spectacular films. Cited on page 151.

123 G. Silva Paiva, A. C. Pavão, E. Alpes de Vasconcelos, O. Mendes & E. F. da Silva, Production of ball-lightning-like luminous balls by electrical discharges in silicon, *Physics Review Letters* **98**, p. 048501, 2007. Cited on page 151.

124 For a recent summary, see S. Parrott, arxiv.org/abs/gr-qc/9711027. See also T. A. Abbott & D. J. Griffiths, Acceleration without radiation, *American Journal of Physics* **53**, pp. 1203–1211, 1985. See also A. Kovetz & G. E. Tauber, Radiation from an accelerated charge and the principle of equivalence, *American Journal of Physics* **37**, pp. 382–385, 1969. Cited on page 152.

125 C. de Almeida & A. Saa, The radiation of a uniformly accelerated charge is beyond the horizon: a simple derivation, *American Journal of Physics* **74**, pp. 154–158, 2006. Cited on page 152.

126 J. Zhang, X. D. Song, Y. C. Li, P. G. Richards, X. L. Sun & F. Waldhauser, Inner core differential motion confirmed by earthquake doublet waveform doublets, *Science* **309**, pp. 1357–1360, 2005. Cited on page 153.

127 An excellent review is E. H. Brandt, Levitation in Physics, *Science* **243**, pp. 349–355, 1989. Cited on pages 154 and 155.

128 See the article by R. Tuckermann, S. Bauerecker & B. Neidhart, Levitation in Ultraschallfeldern – Schwebende Tröpfchen, *Physik in unserer Zeit* **32**, pp. 69–75, February 2001. Liquid drops up to 1 g have been levitated in this way. Cited on page 154.

129 F. C. Moon & P. Z. Chang, *Superconducting Levitation – Applications to Bearings and Magnetic Transportation*, Wiley & Sons, 1994. Cited on pages 154 and 155.

130 W. T. Scott, Who was Earnshaw?, *American Journal of Physics* **27**, pp. 418–419, 1959. Cited on page 154.

131 The trick is to show that div $E = 0$, curl $E = 0$, thus $E\nabla^2 E = 0$ and, from this, $\nabla^2 E^2 \geqslant 0$; there are thus no local electric field maxima in the absence of free charges. The same proof works for the magnetic field. However, bodies with dielectric constants *lower* than their environment *can* be levitated in static electric fields. An example is gas bubbles in liquids, as shown by T. B. Jones & G. W. Bliss, Bubble dielectrophoresis, *Journal of Applied Physics* **48**, pp. 1412–1417, 1977. Cited on page 154.

132 B. Scharlau, V. Nordmeier & H. J. Schlichting, Magnetische Levitation, in Deutsche Physikalische Gesellschaft, (editor) *Didaktik der Physik*, Lehmanns, 2003. Cited on pages 155 and 156.

133 See A. K. Geim, M. D. Simon, M. I. Boamfa & L. O. Heflinger, Magnet levitation at your fingertips, *Nature* **400**, pp. 323–324, 1999. Cited on page 155.

134 The first photographs of a single *ion* were in W. Neuhauser, M. Hohenstatt, P. E. Toschek & H. Dehmelt, Localized visible Ba$^+$ mono-ion oscillator, *Physical Review A* **22**, pp. 1137–1140, 1980. See also D. J. Wineland & W. M. Itano, *Physics Letters A* **82**, p. 75, 1981, as well as F. Dietrich & H. Walter, *Physical Review Letters* **58**, p. 203, 1987.

For single *atoms*, see photographs in Z. Hu & H. J. Kimble, *Optics Letters* **1**, p. 1888, 1994, F. Ruschewitz, D. Bettermann, J. L. Peng & W. Ertmer, *Europhysics Letters* **34**, p. 651, 1996, D. Haubrich, H. Schadwinkel, F. Strauch, B. Ueberholz, R. Wynands & D. Meschede, *Europhysics Letters* **34**, p. 663, 1996. Cited on page 155.

135 See for example Mark Buchanan, And God said...let there be levitating strawberries, flying frogs and humans that hover over Seattle, *New Scientist* pp. 42–43, 26 July 1997, or C. Wu, Floating frogs, *Science News* **152**, pp. 632–363, 6 December 1997, and C. Wu, Molecular magnetism takes off, *Physics World* April 1997, page 28. The experiments by Andre Geim, Jan Kees Maan, Humberto Carmona and Peter Main were made public by P. Rodgers, *Physics World* **10**, p. 28, 1997. Some of the results can be found in M. V. Berry & A. K. Geim, Of flying frogs and levitrons, *European Journal of Physics* **18**, pp. 307–313, 1997. See also their www-hfml.sci.kun.nl/hfml/levitate.html website. Cited on page 155.

136 The well-known toy allows levitation without the use of any energy source and is called the 'Levitron'. It was *not* invented by Bill Hones of Fascination Toys & Gifts in Seattle, as the www.levitron.com website explains. The toy is discussed by Ron Edge, Levitation using only permanent magnets, *Physics Teacher* **33**, p. 252, April 1995. It is also discussed in M. V. Berry, The Levitron^{TM}: an adiabatic trap for spins, *Proceedings of the Royal Society A* **452**, pp. 1207–1220, 1996, (of Berry's phase fame) as well as by M. D. Simon, L. O. Heflinger & S. L. Ridgeway, Spin stabilized magnetic levitation, *American Journal of Physics* **65**, pp. 286–92, 1997, and by T. B. Jones, M. Washizu & R. Gans, Simple theory for the Levitron, *Journal of Applied Physics* **82**, pp. 883–889, 1997. Cited on page 155.

137 The drill trick and the building of a Levitron are described in the beautiful lecture script by Josef Zweck, *Physik im Alltag*, Skript zur Vorlesung im WS 1999/2000 der Universität Regensburg. Cited on page 156.

138 The prediction about quantized levitation is by Stephen B. Haley, Length quantization in levitation of magnetic microparticles by a mesoscopic superconducting ring, *Physical Review Letters* **74**, pp. 3261–3264, 1995. The topic is discussed in more detail in Stephen B. Haley, Magnetic levitation, suspension, and superconductivity: macroscopic and mesoscopic, *Physical Review B* **53**, p. 3506, 1996, reversed in order with Stephen B. Haley, Quantized levitation of superconducting multiple-ring systems, *Physical Review B* **53**, p. 3497, 1996, as well as Stephen B. Haley, Quantized levitation by multiply-connected superconductors, LT-21 Proceedings, in *Czechoslovak Journal of Physics* **46**, p. 2331, 1996. In 1998, there was not yet an experimental confirmation (Stephen Haley, private communication). Cited on page 156.

139 Detailed descriptions of many of these effects can be found in the excellent overview edited by Manfred von Ardenne, Gerhard Musiol & Siegfried Reball, *Effekte der Physik und ihre Anwendungen*, Harri Deutsch, 2004. Cited on page 158.

140 R. Buddakian, K. Weninger, R. A. Hiller & Seth J. Putterman, Picosecond discharges and stick–slip friction at a moving meniscus of mercury in glass, *Nature* **391**, pp. 266–268, 15 January 1998. See also *Science News* **153**, p. 53, 24 January 1998. Cited on page 158.

141 Henk Swagten & Reinder Coehoorn, Magnetische tunneljuncties, *Nederlands tijdschrift voor natuurkunde* **64**, pp. 279–283, November 1998. Cited on page 158.

142 H. Ohno, D. Chiba, F. Matsukura, T. Omiya, E. Abe, T. Dietl, Y. Ohno & K. Ohtani, Electric-field control of ferromagnetism, *Nature* **408**, pp. 944–946, 21-28 December 2000. Cited on page 159.

143 This effect was discovered by G. Rikken, B. van Tiggelen & A. Sparenberg, Lichtverstrooiing in een magneetveld, *Nederlands tijdschrift voor natuurkunde* **63**, pp. 67–70, maart 1998. Cited on page 160.

144 Vitalij Pecharsky & Karl A. Gschneidner, Giant magnetocaloric effect in Gd5(Si2Ge2), *Physical Review Letters* **78**, pp. 4494–4497, 1995, and, from the same authors, Tunable magnetic regenerator alloys with a giant magnetocaloric effect for magnetic refrigeration from ~20 to ~2990 K, *Applied Physics Letters* **70**, p. 3299, 1997. Cited on page 160.

145 J. Weissmüller, R. N. Viswanath, D. Kramer, P. Zimmer, R. Würschum & H. Gleiter, Charge-induced reversible strain in a metal, *Science* **300**, pp. 312–315, 11 April 2003. Cited on page 161.

146 A. Ajdari, Electro-osmosis on inhomogeneously charged surfaces, *Physical Review Letters* **75**, pp. 755–758, 1995. Cited on page 161.

147 M. J. Aitken, *Thermoluminescence Dating*, Academic Press, 1985. The precision of the method is far worse that C14 dating, however, as shown by H. Huppertz, *Thermolumineszenzdatierung: eine methodologische Analyse aufgrund gesicherter Befunde*, Peter Lang Verlag, 2000. Cited on page 162.

148 This effect was discovered by J. N. Huiberts, R. Griessen, J. H. Rector, R. J. Wijngarden, J. P. Dekker, D. G. de Groot & N. J. Koeman, Yttrium and lanthanum hydride films with switchable optical properties, *Nature* **380**, pp. 231–234, 1996. A good introduction is R. Griessen, Schaltbare Spiegel aus Metallhydriden, *Physikalische Blätter* **53**, pp. 1207–1209, 1997. Cited on page 163.

149 See any book on thermostatics, such as Linda Reichl, *A Modern Course in Statistical Physics*, Wiley, 2nd edition, 1998. Cited on page 164.

150 The Sun emits about $4 \cdot 10^{26}$ W from its mass of $2 \cdot 10^{30}$ kg, about 0.2 mW/kg; a person with an average mass of 75 kg emits about 100 W (you can check this in bed at night), i.e., about 500 times more. Cited on page 165.

151 See for example, J. M. Aguirregabiria, A. Hernandez & M. Rivas, Velocity fields inside a conducting sphere near a slowly moving charge, *American Journal of Physics* **62**, pp. 462–466, 1994. Cited on page 165.

152 Philip Cohen, Open wide, this won't hurt a bit, *New Scientist* p. 5, 3 February 1996. Cited on page 165.

153 J. E. Avron, E. Berg, D. Goldsmith & A. Gordon, Is the number of photons a classical invariant?, *European Journal of Physics* **20**, pp. 153–159, 1999. Cited on page 166.

154 This is deduced from the $g - 2$ measurements, as explained in his Nobel Prize talk by Hans Dehmelt, Experiments with an isolated subatomic particle at rest, *Reviews of Modern Physics* **62**, pp. 525–530, 1990, and in Hans Dehmelt, Is the electron a composite particle?, *Hyperfine Interactions* **81**, pp. 1–3, 1993. Cited on page 167.

155 A good and short introduction is the paper F. Rohrlich, The self-force and radiation reaction, *American Journal of Physics* **68**, pp. 1109–1112, 2000. Cited on page 168.

156 Distinguishing between the thought 'yes' and 'no' is already possible with a simple electroencephalogram. For a video demonstration of the differentaiation of concepts using brain imaging techniques, see www.youtube.com/watch?v=JVLu5_hvr8s. Cited on page 168.

157 C. G. Tsagas, Magnetic tension and the geometry of the universe, *Physical Review Letters* **86**, pp. 5421–5424, 2001. An overview of the topic is C. G. Tsagas, Geometrical aspects of cosmic magnetic fields, arxiv.org/abs/gr-qc/0112077. Cited on page 170.

158 A. D. Erlykin & A. W. Wolfendale, The origin of cosmic rays, *European Journal of Physics* **20**, pp. 409–418, 1999, Cited on page 172.

159 See for example the beautiful textbook by STEPHEN C. STEARNS & ROLF F. HOEKSTRA, *Evolution: An Introduction*, Oxford University Press, 2000. For fascinating story of evolution for non-specialists, see RICHARD FORTEY, *Life – An Unauthorized Biography*, Harper Collins, 1997, or also MENNO SCHILTHUIZEN, *Frogs, Flies & Dandelions – the Making of Species*, Oxford University Press/ 2001. See also STEPHEN J. GOULD, *The Panda's thumb*, W.W. Norton & Co., 1980, one of the several interesting and informative books on evolutionary biology by the best writer in the field. An informative overview over the results of evolution, with the many-branched family tree that it produced, is given on the phylogeny.arizona.edu/tree website. About the results of evolution for human beings, see the informative text by K. KUSCH & S. KUSCH, *Der Mensch in Zahlen*, Spektrum Akademischer Verlag, 2nd edn., 2000. The epochal work by CHARLES DARWIN, *On the Origin of Species*, can be found on the web, e.g. on the entisoft.earthlink.net/origspec.htm and on the www.darwin-online.org.uk websites. Cited on page 175.

160 A simple description is MALCOLM ROSS MACDONALD, *The Origin of Johnny*, Jonathan Cape, 1976. See also BAS HARING, *Kaas en de evolutietheorie*, Houtekiet, 2001. Cited on page 175.

161 RICHARD BANDLER, *Using Your Brain for a Change*, Real People Press, p. 18, 1985. Cited on page 175.

162 There is disagreement among experts about the precise timing of this experience. Some say that only birth itself is that moment. However, there are several standard methods to recall memories of early life, even of the time before birth; one is described and used by J. K. STETTBACHER, *Making Sense of Suffering*, Penguin, New York, 1991, translated from the German original. Even more impressive examples are given by N. J. MAYER, *Der Kainkomplex – neue Wege der systemischen Familientherapie*, Integral Verlag, 1998. Cited on page 175.

163 SANJIDA O'CONNELL, *Mindreading – How We Learn to Love and Lie*, Arrow, 1998. This interesting book describes the importance of lying in the development of a human being, and explains the troubles those people have who cannot read other minds and thus cannot lie, such as autists. Cited on pages 177 and 213.

164 The approach to describe observations as related parts is called structuralism; the starting point for this movement was de Saussure's *Cours de linguistique générale* (see the footnote on page 189). A number of thinkers have tried to use the same approach in philosophy, mythology and literature theory, though with little success. An overview of the (modest) success of structuralism in linguistics and its failure in other fields is given by L. JACKSON, *The Poverty of Structuralism: Literature and Structuralist Theory*, Longman, 1991. The author argues that when one reduces systems to interactions, one neglects the specific content and properties of the elements of the system, and this approach prevents a full understanding of the system under discussion. Cited on page 177.

165 For a view of the mental abilities different from that of Piaget (described on page 177), a presently much discussed author is the Soviet experimental psychologist Lev Vigotsky, whose path-breaking ideas and complicated life are described, e.g., in LEV VIGOTSKY, *Mind in Society*, Harvard University Press, 1978, or in RENÉ VAN DER VEER & JAAN VALSINER, *Understanding Vigotsky: a Quest for Synthesis*, Blackwell Publishers, 1994. More extensive material can be found in the extensive work by RENÉ VAN DER VEER & JAAN VALSINGER, *The Vigotsky Reader*, Blackwell, 1994. Cited on page 178.

166 A somewhat unconventional source for more details is the beautiful text by BRUNO BETTELHEIM, The Uses of Enchantment: the Meaning and Importance of Fairy Tales, Knopf, 1976. Cited on page 178.

167 A simple introduction is MANFRED SPITZER, *Lernen – Gehirnforschung und Schule des Lebens*, Elsevier, 2007. Cited on page 178.

168 Quoted in V. HARLEN, R. RAPPMANN & P. SCHATA, *Soziale Plastik – Materialien zu Joseph Beuys*, Achberger Verlag, 1984, p. 61. Cited on page 178.

169 The problems appearing when one loses the ability to classify or to memorise are told in the beautiful book by the neurologist OLIVER SACKS, *The Man Who Mistook His Wife for a Hat*, Picador, 1985, which collects many case studies he encountered in his work. More astonishing cases are collected in his equally impressive text *An Anthropologist on Mars*, Picador, 1995.

   See also the beautiful text DONALD D. HOFFMAN, *Visual Intelligence – How We Create What We See*, W.W. Norton & Co., 1998, and the www.cogsci.uci.edu/~ddhoff website associated to it. Cited on pages 178 and 183.

170 For a passionate introduction to the connections between language and the brain from a Chomskian perspective, see the bestselling book by STEVEN PINKER, *The Language Instinct – How the Mind Creates Language*, Harper Perennial, 1994. The green idea sentence is discussed in a chapter of the book. Cited on pages 178, 218, and 236.

171 An introduction to neurology is JOSEPH LEDOUX, *Synaptic Self: How Our Brains Become Who We Are*, Viking Press, 2002. Cited on page 178.

172 Another good introduction into the study of classifiers is JAMES A. ANDERSON, *An Introduction to Neural Networks*, MIT Press, 1995. An introduction to computer science is given in GLENN BROOKSHEAR, *Computer Science, An Overview*, 6th edition, Addison Wesley, 2000, or in RICK DECKER & STUART HIRSHFIELD, *The Analytical Engine: An Introduction to Computer Science Using the Internet*, Brooks/Cole Publishers, 1998. Cited on page 178.

173 An overview of the status of the research into the origin of bipedalism is given by B. WOOD, Four legs good, two legs better, *Nature* **363**, pp. 587–588, 17 June 1983. Cited on page 180.

174 A good introduction to neural nets is J. HERTZ, A. KROGH & R. PALMER, *Introduction to the Theory of Neural Computation*, Addison Wesley, 1991. Cited on page 181.

175 Quoted from H. EVES, *Mathematical Circles Squared*, Prindle, Weber ans Schmidt, 1972. Cited on page 182.

176 More about the connection between entropy and computers can be found in the classic paper by R. LANDAUER, Irreversibility and heat generation in the computing process, *IBM Journal of Research and Development* **5**, pp. 183–191, 1961, and in C. H. BENNETT & R. LANDAUER, The fundamental physical limits of computation, *Scientific American* **253**, pp. 48–56, 1985. Cited on page 184.

177 W. H. ZUREK, Thermodynamic cost of computation, algorithmic complexity and the information metric, *Nature* **341**, pp. 119–124, 14 August 1989. Cited on page 184.

178 L. SZILARD, Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen, *Zeitschrift für Physik* **53**, p. 840, 1929. This classic paper can also be found in English translation in the collected works by Leo Szilard. Cited on page 184.

179 J.J. Hopfield/ *Nature* **376**, pp. 33–36, 1995. This paper by one of the fathers of the field presents one possibility by which the *timing* of nerve signals, instead of the usually assumed firing frequency, could also carry information. Cited on page 185.

180 The details of the properties of the firing patterns of neurons are nicely described in the article by M. MAHOWALD & R. DOUGLAS, A silicon neuron, *Nature* **354**, pp. 515–518,

19/26 December 1991, in which they show how to simulate a neuron's electrical behaviour using a silicon circuit. Cited on page 185.

181 A. Mechelli, J. T. Crinion, U. Noppeney, J. O'Doberty, J. Ashburner, R. S. Frackowiak & C. J. Price, Neurolinguistics: structural plasticity in the bilingual brain, *Nature* **431**, p. 757, 2004. Cited on page 186.

182 The discussion whether the brain is or is not superior to a computer is nicely summarised by G. Vollmer, Algorithmen, Gehirne, Computer – Was sie können und was sie nicht können, Teil I und Teil II, *Naturwissenschaften* **78**, p. 481, 1991, and **78**, pp. 533–542, 1991. Cited on page 186.

183 T. Seidel, The role of student characteristics in studying micro teaching-leaning environments, *Learning Environments Research* **9**, pp. 253–257, 2006. Cited on page 187.

184 The results with children are due to Niels Birbaumer, those for stage performers to Boris Kleber, both at the Universität Tübingen. More information is found on www.dgbfb.de and on applied-neuroscienc.org. Cited on page 188.

185 J. T. Choi & A. J. Bastian, Adaptation reveals independent control networks for human walking, *Nature Neuroscience* **10**, pp. 1055–1062, 2007. Cited on page 188.

186 For slightly different definitions and a wealth of other interesting information about language, see the beautiful book by David Crystal, *The Cambridge Encyclopedia of Language*, Cambridge University Press, 1987. Cited on page 189.

187 However, the language with the largest available dictionary is Dutch, with the 40 volumes of the Wordenboek der Nederlandsche Taal, which appeared between 1864 and 1998. It has almost 400 000 entries. Cited on page 191.

188 The list and the remark on discovery on concepts is due to a personal communication from Anna Wierzbicka. A longer list is published in her book *Semantics, Primes and Universals*, Oxford University Press, 1996. Cited on pages 192 and 211.

189 W. S. Hatcher, *Foundations of Mathematics*, W.B. Saunders Co., Philadelphia, 1968. There is also the article by P. J. Cohen & R. Hersch, Non-Cantorian set theory, *Scientific American* **217**, pp. 104–116, 1967. Cohen was the mathematician who in 1963 proved that the negation of the continuum hypothesis could be added to the axioms of set theory and still produce a consistent theory; he calls such sets *non-Cantorian*. Cited on page 196.

190 See the beautiful article by I. Stewart, Fair shares for all, *New Scientist*, pp. 42–46, 17 June 1995. Cited on page 196.

191 The proof of the independence of the continuum hypothesis came in two parts. First, Kurt Gödel proved in 1940 that an axiom can be consistently added to ZFC set theory so that the continuum hypothesis is correct. Then, in 1963, Paul Cohen proved that an axiom can be consistently added to ZFC set theory so that the continuum hypothesis is false. Cited on page 198.

192 Rudy Rucker, *Infinity and the Mind – the Science and Philosophy of the Infinite*, Bantam, Toronto, 1983. Cited on page 198.

193 This general division of mathematics is nicely explained in the text by Pierre Basieux, *Die Architektur der Mathematik – Denken in Strukturen*, Rororo, 2000. Cited on page 200.

194 Umberto Pelizzari, *L'Homme et la mer*, Flammarion, 1994. Cited on page 200.

195 The issue is treated in Thomas Aquinas, *Summa Theologica*, in question 52 of the first part. The complete text, several thousand pages, can be found on the www.newadvent.org website. We come back to it in the part on quantum theory, in the section on the Pauli exclusion principle. It seems that the whole question goes back to Peter (the) Lombard, *Liber Sententiarum c.* 1150. Cited on page 201.

196 B. C. GALLIVAN, How to fold paper in half twelve times: an "impossible challenge" solved and explained, Histrical Society of Pomona Valley, 2002, also found at www.osb.net/Pomona/12times.htm. See also www.sciencenews.org/20040124/mathtrek.asp. Cited on page 201.

197 I. STEWART, Daisy, daisy, give me your answer, do, *Scientific American*, pp. 76–79, January 1995. This pedagogical article explains how the growth of plants usually leads to flowers whose number of petals is from the Fibonacci series 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, etc. (The figure on page 198 is an example.) Deviations from this ideal case are also explained. The original work are two articles by S. DOUADY & Y. COUDER, La physique des spirales végétales, *La Recherche* **24**, pp. 26–36, 1993, and Phyllotaxis as a self-organized growth process, in *Growth Patterns in Physical Sciences and Biology*, edited by J.M. GARCIA-RUIZ & al., Plenum Press, 1993. Cited on pages 201 and 208.

198 H. DAVSON, *The Eye*, Academic Press, 1962. Cited on page 201.

199 See the akbar.marlboro.edu/~mahoney/cube/NxN.txt website. Cited on page 202.

200 An introduction to the surreal numbers is given by the article by POLLY SHULMAN, Infinity plus one, and other surreal numbers, *Discover*, pp. 99–105, December 1995. There is also the text by D. KNUTH, *Surreal Numbers: How two ex-Students Turned on to Pure Mathematics and Found Total Happiness*, Addison Wesley, 1974, or www-cs-faculty.stanford.edu/~knuth/sn.html. The usually quoted references on the topic include JOHN H. CONWAY, *On Numbers and Games*, Academic Press, 1976, E. R. BERLEKAMP, J. H. CONWAY & R. K. GUY, *Winning Ways for Your Mathematical Plays, Volume I: Games in General*, Academic Press, 1982, and H. GONSHOR, *An Introduction to Surreal Numbers*, Cambridge University Press, 1986. Cited on pages 202 and 204.

201 This beautiful problem is discussed by IAN STEWART, A bundling fool beats the wrap, *Scientific American,* pp. 109–111, June 1993. In four dimensions, the answer is known to lie somewhere between 50 000 and 100 000, whereas the five-dimensional answer is conjectured to be 'never'. Cited on page 205.

202 A. PAIS, *Niels Bohr's Times: in Physics, Philosophy and Polity*, Oxford University Press, 1991, page 176. Cited on page 205.

203 EUGENE WIGNER, *Symmetries and Reflections*, Indiana University Press, 1962. Cited on page 205.

204 ALFRED TARSKI, *Introduction to Modern Logic*, Dover, 1946. See also the famous children's book by the mathematician and photographer LEWIS CARROLL, *Alice in Wonderland*. Cited on page 206.

205 Gˆran WIKELL, The layout of digits on pushbutton telephones – a review of the literature, *Tele* **34**, pp. 34–42, 1982. Cited on page 207.

206 A clear overview of philosophy of science, often called *epistemology*, without unnecessary detail, is given by ROBERT BLANCHÉ, L'Epistémologie, Presses Universitaires de France, 1972. Cited on page 209.

207 About the different aspects of falsifiability of general statements it is commonplace to cite the work by the epitemologist Karl Popper (1902–1994), especially his long and boring book *Logik der Forschung*, first published in 1934. The reason for this boredom is that Popper's work is simply a warming-up of Pierre Duhem's ideas. Cited on page 214.

208 For a good way of making blood that liquefies, see L. GARLASCHELLI, F. RAMACCINI & S. DELLA SCALA, Working bloody miracles, *Nature* **353**, p. 507, 1991. The *Grand dictionnaire universel du XIXe siècle*, by PIERRE LAROUSSE, also contains a recipe; it was again shown to the public in the 1980s by Henri Broch. A wonderful and classic text is

HARRY HOUDINI, *Miracle Mongers and their Methods*, Prometheus Books, Buffalo, 1981. The original, written in 1920, by the world famous magician named 'The Great Houdini', is also available on the etext.lib.virginia.edu/toc/modeng/public/HouMirM.html website. The milk drinking Indian statues were common around the world in 1994 and 1995. About healers, see JAMES RANDI, *Flim-flam!*, Prometheus Books, Buffalo, New York, 1987, and the exquisite book by HANS CONRAD ZANDER, *Warum ich Jesus nicht leiden kann*, Rowohlt, 1994. Cited on page 215.

209 ROGER PENROSE, *The Road to Reality: A Complete Guide to the Laws of the Universe*, Jonathan Cape, 2004, page 378. Cited on page 233.

210 JOHN HORGAN, *The End of Science – Facing the Limits of Knowledge in the Twilight of the Scientific Age*, Broadway Books, 1997, pp. 31–32, and chapter 2, note 2. Cited on pages 217 and 233.

211 For an opinion completely contrary to the one described here, see the book by GREGORY J. CHAITIN, *The Limits of Mathematics*, Springer Verlag, 1997, which can also be found on the author's website at www.cs.auckland.ac.nz/CDMTCS/chaitin/lm.html, along with his other works. Chaitin has devoted most of his life to the questions discussed in the section, especially on computability. Cited on page 218.

212 See the book by J. BARWISE & J. ETCHEMENDY, *The Liar*, Oxford University Press, New York, 1987. Cited on page 218.

213 DEMOSTHENES, *Third Olynthiac*, section 19. Cited on page 219.

214 This definition (statement 4.11) and many other statements about science are in the beautiful and rightly famous text by LUDWIG WITTGENSTEIN, *Tractatus logico-philosophicus*, Edition Suhrkamp, 1963. It gives a condensed summary of the basis of science, thought and language in a collection of highly structured and numbered sentences. Cited on pages 222 and 227.

215 See M. DRESDEN, The Klopsteg memorial lecture, *American Journal of Physics* **66**, pp. 468–482, 1998. Cited on page 223.

216 Well-known books are e.g. FRIEDRICH KOHLRAUSCH, *Praktische Physik*, Teubner, 24. Auflage, 1996. Cited on page 223.

217 Results are described e.g. in L. BERGMANN & C. SCHÄFER, *Lehrbuch der Experimentalphysik*, Band I, II, III und IV, W. de Gruyter. Cited on page 223.

218 LANDOLT-BÖRNSTEIN, edited by K. -H. HELLWEGE & O. MADELUNG, *Zahlenwerte und Funktionen aus Naturwissenschaften und Technik*, Neue Serie, Springer Verlag, Berlin, 1984. This series of more than one hundred volumes contains all important observations in the domain of physics. Cited on page 223.

219 The origin of this incorrect attribution is the book by GERHARD SZCZESNY, *Brecht, Leben des Galilei – Dichtung und Wirklichkeit*, Ullstein, Berlin 1966, p. 6. The statement has never been made by Galilei; this issue has been discussed at length in specialist circles, e.g. by F. KLEINERT, "Messen was meßbar ist" - Über ein angebliches Galilei-Zitat, *Berichte zur Wissenschaftgeschichte* **11**, p. 221, 1988, or on the internet by Peter Jaencke. Cited on page 224.

220 The strange and sometimes dangerous consequences of beliefs can be found e.g. in MARTIN GARDNER, *Fads and Fallacies*, Dover, 1957, and in JAMES RANDI, *Faith Healers*, Prometheus Books, 1989. The million dollar prize for showing any paranormal or supernormal effect is available from his www.randi.org website. Cited on page 228.

221 See the nice collection of cranks on the www.crank.net website. Cited on page 228.

222 The opposite view on the emergence of properties is strongly defended in the book by Robert Laughlin, *A Different Universe: Reinventing Physics from the Botton Down* Basic Books, 2005, or by P. Jensen, Particle physics and our everyday world, *Physics Today* pp. 58–59, July 1998. Cited on page 229.

223 See page 133 of the bibliography by John Bowlby, *Charles Darwin*, Pimlico, 1991. Cited on page 230.

224 A beautiful introduction to Greek philosophy is Eduard Zeller, *Outlines of the History of Greek Philosophy*, Dover, 1980, a reprint of a book published in 1928. Among others, it gives a clear exposition of the philosophy of Democritus and the other presocratics. Cited on page 234.

225 The famous quote is found at the beginning of chapter XI, 'The Physical Universe', in Arthur Eddington, *The Philosophy of Physical Science*, Cambridge, 1939. Cited on page 234.

226 Giuseppe Fumagalli, *Chi l'ha detto?*, Hoepli, Milano, 1983. Cited on page 236.

227 See Jean-Paul Dumont, *Les écoles présocratiques*, Folio Essais, Gallimard, p. 653, 1991. Cited on page 236.

228 For a beautiful text on fractals, see the footnote on page 51. Cited on page 238.

229 As has been pointed out by René Descartes. Cited on page 240.

230 The famous carbon ^{12}C resonance was found by Willy Fowler, as described in Margaret Burbridge, G. R. Burbridge, W. A. Fowler & F. Hoyle, Synthesis of the elements in stars, *Reviews of Modern Physics* **29**, pp. 547–560, 1957. Cited on page 243.

231 An extensive overview of the topic is given in the thick book by John D. Barrow & Frank J. Tipler, *The Anthropic Cosmological Principle*, Oxford University Press, 1986. The term itself is due to Brandon Carter, who coined it in 1973 and presented it in a symposium devoted to the 500th anniversary of Nicolaus Copernicus. For more literature, see Yuri I. Balashov, Resource Letter AP-1: the anthropic principle, *American Journal of Physics* **59**, pp. 1069–1076, 1991. Cited on page 243.

232 Voltaire, *Candide ou l'optimisme*, 1759. See the footnote on page 206. The book is so good that it was still being seized by the US customs in 1930, and the US post office refused to transport it as late as 1944. For more details, search for 'banned books online' on the world-wide web. Cited on page 244.

233 The number of books on consciousness is large and the contents not always interesting, and often not based on fact, as shown by Karl R. Popper & John Eccles, *The Self and its Brain – an Argument for Interactionism*, Rutledge, 1993. Cited on page 244.

234 See e.g. the *Encyclopedia Britannica*, Macropaedia, in the entry on animal behaviour. Cited on page 245.

235 A straight and informative introduction to the work and ideas of Joseph Beuys (in German) is by Renate Georgi, Joseph Beuys, RAAbits Kunst, Raabe Fachverlag, September 1996. Cited on page 245.

236 Two studies, one by R.P. Ebstein & al., Dopamine D4 receptor (D4DR) exon III polymorphism associated with human personality trait of novelty seeking, *Nature Genetics* **12**, pp. 78–80, January 1996, and another study by J. Benjamin & al., Population and familial association between the D4 dopamine receptor gene and measures of novelty seeking, *Nature Genetics* **12**, pp. 81–84, January 1996, found that people with a special form of the D4 dopamine receptor gene, or D4DR, are more prone to novelty seeking than people with the usual form. The D4DR gene regulates the formation of dopamine receptors, a chemical

messenger in the brain that has been a candidate for some time for a substance involved in the propensity for novelty seeking. Cited on page 246.

237 See Jacques Hadamard, *The Mathematician's Mind – The Psychology of Invention in the Mathematical Field*, Princeton Science Library, 1996. For a modern perspective, see Pierre de Gennes, *Fragile Objects: Soft Matter, Hard Science, and the Thrill of Discovery*, Springer, 1996. where de Gennes criticizes certain educational systems that put too much stress on mathematics, thus destroying creativity. Cited on page 246.

238 Voltaire writes this in his *Catalogue pour la plupart des écrivains français qui ont paru dans Le Siècle de Louis XIV, pour servir à l'histoire littéraire de ce temps* (1752). Cited on page 247.

239 This is from the beautiful booklet by Bert Hellinger, *Verdichtetes*, Carl-Auer Systeme Verlag, 1996. Cited on page 247.

240 For example, one needs the courage to face envy. About this topic see the classic text by Helmut Schoeck, *Der Neid*, 1966, published in English as *Envy: A Theory of Social Behavior*, 1969. It is the standard work in the field. Cited on page 247.

241 Bill McGuire, *A Guide to the End of the World: Everything You Never Wanted to Know*, Oxford University Press, 2002. On past disasters, see introduction by Tony Hallam, *Catastrophes and Lesser Calamities – the Causes of Mass Extinctions*, Oxford University Press, 2004. Cited on page 251.

242 Le Système International d'Unités, Bureau International des Poids et Mesures, Pavillon de Breteuil, Parc de Saint Cloud, 92310 Sèvres, France. All new developments concerning SI units are published in the journal *Metrologia*, edited by the same body. Showing the slow pace of an old institution, the BIPM launched a website only in 1998; it is now reachable at www.bipm.fr. See also the www.utc.fr/~tthomass/Themes/Unites/index.html website; this includes the biographies of people who gave their names to various units. The site of its British equivalent, www.npl.co.uk/npl/reference, is much better; it provides many details as well as the English-language version of the SI unit definitions. Cited on page 256.

243 The bible in the field of time measurement is the two-volume work by J. Vanier & C. Audoin, *The Quantum Physics of Atomic Frequency Standards*, Adam Hilge, 1989. A popular account is Tony Jones, *Splitting the Second*, Institute of Physics Publishing, 2000.
       The site opdaf1.obspm.fr/www/lexique.html gives a glossary of terms used in the field. For precision *length* measurements, the tools of choice are special lasers, such as mode-locked lasers and frequency combs. There is a huge literature on these topics. Equally large is the literature on precision *electric current* measurements; there is a race going on for the best way to do this: counting charges or measuring magnetic forces. The issue is still open. On *mass* and atomic mass measurements, see Volume II, page 64. On high-precision *temperature* measurements, see page 425. Cited on page 257.

244 The unofficial prefixes were first proposed in the 1990s by Jeff K. Aronson of the University of Oxford, and might come into general usage in the future. Cited on page 258.

245 The various concepts are even the topic of a separate international standard, ISO 5725, with the title *Accuracy and precision of measurement methods and results*. A good introduction is John R. Taylor, *An Introduction to Error Analysis: the Study of Uncertainties in Physical Measurements*, 2nd edition, University Science Books, Sausalito, 1997. Cited on page 259.

246 P. J. Mohr & B. N. Taylor, CODATA recommended values of the fundamental physical constants: 1998, *Reviews of Modern Physics* 59, p. 351, 2000. This is the set of constants resulting from an international adjustment and recommended for international use by the Committee on Data for Science and Technology (CODATA), a body in the International Council of Scientific Unions, which brings together the International Union of Pure and Applied

Physics (IUPAP), the International Union of Pure and Applied Chemistry (IUPAC) and other organizations. The website of IUPAC is www.iupac.org. Cited on pages 260 and 261.

247 Some of the stories can be found in the text by N. W. WISE, *The Values of Precision*, Princeton University Press, 1994. The field of high-precision measurements, from which the results on these pages stem, is a world on its own. A beautiful introduction to it is J. D. FAIRBANKS, B. S. DEAVER, C. W. EVERITT & P. F. MICHAELSON, eds., *Near Zero: Frontiers of Physics*, Freeman, 1988. Cited on page 260.

248 The details are given in the well-known astronomical reference, KENNETH SEIDELMANN, *Explanatory Supplement to the Astronomical Almanac*, 1992. Cited on page 264.

249 For information about the number π, and about some other mathematical constants, the website oldweb.cecm.sfu.ca/pi/pi.html provides the most extensive information and references. It also has a link to the many other sites on the topic, including the overview at mathworld.wolfram.com/Pi.html. Simple formulae for π are

$$\pi + 3 = \sum_{n=1}^{\infty} \frac{n\,2^n}{\binom{2n}{n}} \tag{113}$$

or the beautiful formula discovered in 1996 by Bailey, Borwein and Plouffe

$$\pi = \sum_{n=0}^{\infty} \frac{1}{16^n} \left(\frac{4}{8n+1} - \frac{2}{8n+4} - \frac{1}{8n+5} - \frac{1}{8n+6} \right) . \tag{114}$$

The mentioned site also explains the newly discovered methods for calculating specific binary digits of π without having to calculate all the preceding ones. The known digits of π pass all tests of randomness, as the mathworld.wolfram.com/PiDigits.html website explains. However, this property, called *normality*, has never been proven; it is the biggest open question about π. It is possible that the theory of chaotic dynamics will lead to a solution of this puzzle in the coming years.

Another method to calculate π and other constants was discovered and published by D. V. CHUDNOVSKY & G. V. CHUDNOVSKY, The computation of classical constants, *Proceedings of the National Academy of Sciences (USA)* **86**, pp. 8178–8182, 1989. The Chudnowsky brothers have built a supercomputer in Gregory's apartment for about 70 000 euros, and for many years held the record for calculating the largest number of digits of π. They have battled for decades with Kanada Yasumasa, who held the record in 2000, calculated on an industrial supercomputer. However, the record number of (consecutive) digits in 2010 was calculated in 123 days on a simple desktop PC by Fabrice Bellard, using a Chudnovsky formula. Bellard calculated over 2.7 million million digits, as told on bellard.org. New formulae to calculate π are still occasionally discovered.

For the calculation of Euler's constant γ see also D. W. DeTEMPLE, A quicker convergence to Euler's constant, *The Mathematical Intelligencer*, pp. 468–470, May 1993.

Note that little is known about the basic properties of some numbers; for example, it is still not known whether $\pi + e$ is a rational number or not! (It is believed that it is not.) Cited on page 266.

Challenge 295 r

CREDITS

cent Darley, Johan Linde, Joseph Hertzlinger, Rick Zaccone, John Warkentin, Ulrich Diez, Uwe Siart, Will Robertson, Joseph Wright Enrico Gregorio, Rolf Niepraschk and Alexander Grahn.

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.

All illustrations and animations in the text were made available by the copyright holders. A warm thank you to all of them. They are mentioned in the image and film credit sections. In particular, Lucas Barbosa and José Antonio Díaz Navas produced special animations for this text and Luca Gastaldi, Antonio Martos and Ulrich Kolberg produced special images for this text. The typesetting and book design is due to the professional consulting of Ulrich Dirr. The typography was improved with the help of Johannes Küster. The design of the book and its website owe also much to the suggestions and support of my wife Britta.

Since May 2007, the electronic edition and distribution of the Motion Mountain text is generously supported by the Klaus Tschira Foundation.

FILM CREDITS

The animations of a plane electromagnetic wave on page 85 are copyright and courtesy by Thomas Weiland and taken from on his website www.temf.de at the Technische Universität Darmstadt. The animation of the electromagnetic field emitted by an oscillating charge on page 98 is copyright and courtesy by Daniel Schroeder. He will post it on his website physics.weber.edu/schroeder/mrr/MRRtalk.html one day. The animation of the electromagnetic field emitted by an oscillating dipole on page 99 is copyright and courtesy by Daniel Weiskopf and can be found on his website www.vis.uni-stuttgart.de/~weiskopf.

IMAGE CREDITS

The photograph of a peak in the Himalaya on the front cover is courtesy and copyright by Dave Thompson and used to be on his website www.daveontrek.co.uk. The photograph of a circular rainbow on page 14 is copyright and courtesy of Oat Vaiyaboon, and taken from his flickr collection; his website is hangingpixels.com. The photographs of electric field lines on page 18 are copyright and courtesy of Eli Sidman, from the Technical Service Group of the Massachusetts Institute of Technology, and found on the group website on tsgphysics.mit.edu. The photograph of a rubbed comb's effect on water on page 16 is copyright and courtesy of Robert Fritzius and found on his website www.datasync.com/~rsf1/fun/bend-w.htm. The ground-breaking graphics of electric fields on page 18 are copyright and courtesy of the TEAL group at MIT, and found on their website web.mit.edu/8.02t/www/802TEAL3D/visualizations/guidedtour/Tour.htm. The photograph of lightning on page 19 is copyright Steven Horsburgh (see www.horsburgh.com) and used with his permission. The photograph of the Kelvin generator on page 20 is courtesy and copyright of Harald Chmela and taken from his website www.hcrs.at. The photograph of a white shark on page 22 is copyright and courtesy of Klaus Jost and found on his beautiful website at www.jostimages.com. The pictures of the charge conservation experiment on page 21 and of a red hot oven on page 120 are copyright and courtesy of Wolfgang Rueckner. The photograph of *M. bavaricum* on page 37 is copyright by Marianne Hanzlik and is courtesy of Nicolai Petersen. The pictures of the Tesla coil on page 49 are courtesy and copyright of Robert Billon, and found on his website f3wm.free.fr. The photograph of the electrified hair on the playground, on page 50, is courtesy and copyright of Evan Keller and found on his website www.flickr.com/photos/evanrkeller. The photograph of a lifter on page 52 is courtesy and copyright of Jean-Louis Naudin; more information can be found on his website www.jlnlabs.org. The ocean figure on page 53 is courtesy of Stefan Maus, and taken from his www.gfz-potsdam.de/pb2/pb23/SatMag/

ocean_tides.html website. The simple motor photograph on page 54 is courtesy and copyright of Stefan Kluge. The comic on page 60 is copyright and courtesy of Randall Munroe, and found on his website xkcd.com. The Gauss rifle on page 51 is courtesy and copyright Simon Quellen Field and found on his website www.sci-toys.com. The photograph of the floating bed model on page 57 and the computer graphics of the imagined full-size floating bed are courtesy and copyright Janjaap Ruissenaars at www.UniverseArchitecture.com. The prism photograph on pages 84 and on 104 is by Susan Schwartzenberg and courtesy and copyright of the Exploratorium, found at www.exploratorium.edu. The picture of the rainbow on page 76 is from the NOAA website. The photographs of the beautifully simple remote control shown on page 87 are copyright and courtesy of Guido Pegna, and found on his website www.pegna.com. The secondary rainbow picture on page 88 is courtesy and copyright of Antonio Martos. The supernumerary rainbow picture on page 88 is courtesy and copyright of Wolfgang Hinz and from his website www.meteoros.de. The interference figure on page 89 is copyright and courtesy of Rüdiger Paschotta and found on his free laser encyclopedia at www.rp-photonics.com. The combined infrared and visible rainbow picture on page 90 is courtesy and copyright of Stefan Zeiger collection at www.photo.net/photodb/member-photos?include=all&user_id=439012. The photographs of birefringence on page 91 are copyright and courtesy of Roger Weller, from his website skywalker.cochise.edu/wellerr/mineral/calcite/calcite1.htm, Brad Amos, from his website homepage.ntlworld.com/w.amos2/BradAmos'sWebsite, and Martin Pietralla, from his lecture material. The images formed by lenses on page 128 are copyright and courtesy of Eric Kirchner and found in his paper cited in the text. The photo of Poisson's spot on page 133 is courtesy and copyright of Christopher Jones, and taken from his website www.union.edu/PUBLIC/PHYDEPT/jonesc/scientific_photos.htm. The photograph of a superior mirage on page 126 is courtesy and copyright by Thomas Hogan and found on his website home.centurytel.net/Arkcite/looming.htm. The photograph of an inferior mirage on page 126 is courtesy and copyright by Andy Barson and found on his website www.andybarson.co.uk. The solar furnace photograph on page 120 is courtesy and copyright of Gerhard Weinrebe. The photo of comet McNaught on page 101 is courtesy and copyright of Flagstaffotos. The photograph of the solar green flash on page 106 is copyright and courtesy of Andrew Young and part of his extensive and fascinating website at mintaka.sdsu.edu/GF; the lunar green flash photograph is copyright and courtesy of Laurent Laveder and taken from his beautiful site at www.PixHeaven.net. The picture of milky water on page 106 was made for this text and is copyright by Antonio Martos. The fogbow photograph on page 108 is courtesy and copyright of Michel Tournay, and can be found on his website www.spacew.com/gallery/Micheltournay. The photograph of a split rainbow on page 108 is courtesy and copyright of Eva Seidenfaden, and can be found on her website www.paraselene.de. The photograph of the sixfold rainbow on page 108 is courtesy and copyright of Terje Nordvik, and can be found on antwrp.gsfc.nasa.gov/apod/ap070912.html. The photograph of the red rainbow on page 108 is courtesy and copyright of Zhu XiaoJin, and can be found on his collection at www.cs.cmu.edu/~zhuxj/astro/. The photograph of the moon rainbow on page 108 is courtesy and copyright of Laurent Laveder, and can be found on his collection at www.pixheaven.net. The photograph of parhelia on page 109 is courtesy and copyright of Phil Appleton and found on the website www.astronet.ru/db/xware/msg/1174325/solsticehalo_appleton.jpg.html. The infrared photograph on page 139 is copyright and courtesy of Serge Augustin. The X-ray tomographs on page 137 are copyright and courtesy of Manuel Dierick and his research group at the University of Ghent and found on his website at www.ugct.ugent.be. The photograph of the sunflower on page 140 is copyright and courtesy of Andrew Davidhazy and found on his website www.rit.edu/~andpph. The pictures of retinas on page 141 are courtesy and copyright of Austin Roorda. The microscope picture on page 134 is copyright and courtesy of Stefan Hell. The pictures showing colour blindness on page 144 are courtesy and copyright of Michael Douma, from his splendid website at webexhibits.org/causesofcolor/2.

html. The photographs of Ames rooms on page 145 are courtesy and copyright of Sergio Davini, taken from his website www.flickr.com/photos/mosso, and courtesy and copyright of David Darling, taken from his encyclopedic website www.daviddarling.info. The cloud photographs on page 148 are courtesy of NASA. The pictures of laboratory plasma clouds that resemble ball lightning on page 151 are courtesy and copyright by Sergei Emelin and Alexei Pirozerski and taken from their website balllightning.narod.ru. The drawings of the interior of the Earth on page 153 are copyright and courtesy of MPI-Chemie, Mainz/GEO and can be found in the brochure at www.mpch-mainz.mpg.de/mpg/deutsch/Panels_B.pdf. They were kindly provided by Mirjana Kotowski. The photographs of diamagnetic levitation on page page 156 are copyright and courtesy of Joachim Schlichting and can be found on his website www.uni-muenster.de/Physik/DP. The photographs of the levitation of a rotating sphere on page page 157 are copyright and courtesy of Kay Kublenz and can be found on his website www.schwebemagnet.de. The photograph of Fibonacci washers on page page 208 is copyright and courtesy of Donald Simanek and can be found on his website www.lhup.edu/~dsimanek. The historical portraits of physicists in the text do not have copyright, except where mentioned. 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.terra.vzw.org. All drawings are copyright by Christoph Schiller. If you suspect that your copyright is not correctly given or obtained, this has not done on purpose; please contact me in this case.

NAME INDEX

B

Boone

S
Swagten

W

Wu

SUBJECT INDEX

B

BELIEFS

D

DUAL

H

HEAT

L

LASER

M

MEISSNER

S

SUBJECTS

U

UNIFICATION

MOTION MOUNTAIN
The Adventure of Physics – Vol. III
Light, Charges and Brains

Why do change and motion exist?
How does a rainbow form?
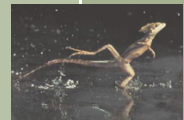What is the most fantastic voyage possible?
Is 'empty space' really empty?
How can one levitate things?
At what distance between two points does it become
    impossible to find room for a third one in between?
What does 'quantum' mean?
Which problems in physics are unsolved?

Answering these and other questions on motion,
this series gives an entertaining and mind-twisting
introduction into modern physics – one that is
surprising and challenging on every page.
    Starting from everyday life, the adventure provides
an overview of the recent results in mechanics,
thermodynamics, electrodynamics, relativity,
quantum theory, quantum gravity and unification.
It is written for undergraduate students and for
anybody interested in physics.

Christoph Schiller, PhD Université Libre de Bruxelles,
is a physicist with more than 25 years of experience
in the presentation of physical topics.