Helmut Moritz

Science, Mind and the Universe

An Introduction to Natural Philosophy

WICHMANN

Alle in diesem Buch enthaltenen Angaben, Daten, Ergebnisse etc. wurden vom Autor nach bestem Wissen erstellt und von ihm und dem Verlag mit größtmöglicher Sorgfalt überprüft. Gleichwohl sind inhaltliche Fehler nicht vollständig auszuschließen. Daher erfolgen die Angaben etc. ohne jegliche Verpflichtung oder Garantie des Verlags oder des Autors. Sie übernehmen deshalb keinerlei Verantwortung und Haftung für etwaige inhaltliche Unrichtigkeiten.

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Moritz, Helmut:
Science, Mind and the Universe : An introduction to natural philosophy /
Helmut Moritz. – Heidelberg : Wichmann, 1995
ISBN 3-87907-274-4

Druck und Verarbeitung: E. Lokay, Reinheim
Gedruckt auf chlorfrei gebleichtem, säurefreiem Papier

ISBN 3-87907-274-4

To my colleagues from

IAG
IUGG
ICSU
ÖAW

in gratitude

Preface

Schuster, bleib bei deinem Leisten!
German proverb

"The rabbi spoke three times. The first talk was brilliant: clear and simple. I understood every word. The second was even better: deep and subtle. I didn't understand much, but the rabbi understood all of it. The third was by far the finest: a great and unforgettable experience. I understood nothing, and the rabbi didn't understand much either." This was one of Niels Bohr's favorite anecdotes (Folse, 1985, p. 258).

All books on philosophy belong to one of the three types of the rabbi's talks. In an introductory text like the present book, the reader is entitled to expect that it belongs to "Rabbi Type 1". This is at least what we have tried to achieve.

Our higher polytechnical schools have frequently become "Technical Universities" or "Universities of Technology". This implies that they intend not only to give a profound scientific or professional education, but also to offer a touch of "universality". Now, the common interdisciplinary background of all scientific, engineering and medical disciplines is becoming increasingly "philosophical".

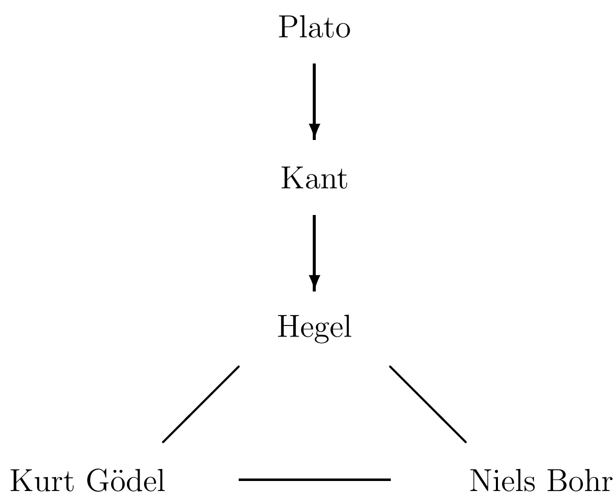
On all aspects of philosophy of mathematics and natural science there are excellent and even brilliant monographs. The intention of the present book is much more modest: to provide an absolutely introductory yet rather systematic and comprehensive textbook taylored to the interests of students of science, technology, and medicine.

This book is written not by a professional philosopher, but by a practicing scientist. The lack of philosophical depth is thus perhaps partly made up by a knowledge of the mentality and interests of students of scientific disciplines and by active research experience in science, in which interdisciplinary and philosophical questions are arising naturally and to an ever increasing extent.

The author's philosophical interests date back to the age of 17 years, when he happened to come across the brilliant book "An Outline of Philosophy" by Bertrand Russell (Russell 1927). Later his interests changed from Russell to his more profound and more many–sided colleague Alfred North Whitehead. To me there is only one modern philosopher of comparable depth and universality: Carl Friedrich von Weizsäcker.

In 1974 and 1975 I gave a course on the philosophy of mathematics and natural sciences at the Graz University of Technology. Later my time was absorbed by other duties but my interest remained, and finally I decided to resume this course again and to write the present book.

The scheme of the book may be illustrated by the following diagram showing the "sign posts" of it in a somewhat provocative way:



To avoid misunderstandings: the names stand for certain directions of thinking and have no direct relation to "greatness". (Otherwise one might ask: why Plato and not Aristotle, why Bohr and not Einstein or Heisenberg?)

Only the first three are "professional" philosophers. Gödel is the greatest of modern logicians, and Bohr, through his "principle of complementarity", is the physicist who has made *dialectics* respectable in the philosophy of natural science. When he was awarded a title of nobility for his scientific merits, he chose for his coat of arms the inscription "*Contraria sunt complementa*".

Dialectics has frequently fallen into discredit by its dogmatic use and misuse, and also Hegel has suffered from this circumstance. Nevertheless dialectic thinking does have its use in natural philosophy, and the justly famous bestseller "Gödel, Escher, Bach" by D.R. Hofstadter is thoroughly dialectical although here dialectics is discreetly disguised under the term of "paradoxes of self–reference".

It is also remarkable that it is rare indeed that *all* three names forming the bottom triangle of the figure are quoted *simultaneously* in the literature: philosophers speak of Hegel, mathematical logicians quote Gödel, and physicists refer to Bohr. Practically the only exception are recent books by C.F. von Weizsäcker and, on a similarly high level, the work "*Filozofija znanosti i humanizam*" by Ivan Supek (SNL, Zagreb, 1991).

Emphasis is also on a basic element of *uncertainty*, which seems to pervade nature and our knowledge of it. It is expressed by Gödel's incompleteness theorem, by fuzzy logic, by Heisenberg's uncertainty relation, by other random fluctuations and random measuring errors, etc., which have fascinated the imagination of mathematicians, physicists, astronomers, and geodesists since C.F. Gauss.

The main intention of the book is didactic simplicity. Another purpose is a more or less systematic unification of trends of philosophy of science in Anglo–Saxon countries and in Central Europe. I believe that not only Kant, but also some other German philosophers (among them Fichte, Hegel, and Weizsäcker) have made contributions to our field which deserve being known also in English–speaking countries. This is one of the several reasons for writing this text in English.

Standard books on the philosophy of science (in German: Wissenschaftstheorie) such as (Carnap 1966), (Hempel 1966) or (Popper 1977) chiefly deal with the logical and methodological basis common to all natural sciences, such as axiomatics and induction. The present text treats such problems also, but more briefly. Emphasis is on general results of modern science (e.g. quantum theory), their interpretation and their implications for standard philosophical problems, such as the relation between matter and mind, the meaning of a law of nature, free will etc.

The term "natural philosophy" has a proud English tradition: physicists from Isaac Newton to W. Thomson (Lord Kelvin) used it for what we would today call "theoretical physics with a philosophical touch". The word "Naturphilosophie" was used in Germany, by Goethe, Schelling, and Hegel, in quite a different, philosophical–speculative sense. Again, the present work seeks to reflect something of both connotations. Here, "natural philosophy" is used in very much the same sense as in Whitrow's well–known book "The Natural Philosophy of Time" (Whitrow 1980).

The book should be comprehensible not only to undergraduate students of mathematics, physics, natural science, technology, and medicine, but also to a broader group of readers interested in philosophical implications of modern science. Mathematical formulas are kept to a minimum: while possibly of help to the mathematically versed reader, they can be bypassed without serious problems. Similarly, the references to selected additional works may provide further information; they should, however, not be necessary to a basic understanding of the present book, which is essentially self–contained.

For didactic purposes even a rigorous logical structure has sometimes been sacrificed: important concepts such as materialism and idealism have been discussed, to increasing detail and on increasing level, in several places of the book. Generally, redundancies and repetitions are not avoided if they improve the readability. In this way, the book resembles an informal university lecture.

Since this is an introductory book, it cannot be very profound. It probably also contains quite a few errors. I have, however, tried to make it interesting reading: it would be a crime to treat such a fascinating subject, one of the greatest adventures of the human spirit, in a dull manner.

Acknowledgments. Writing such a general book requires, but also finds, much more interest and encouragement than a technical monograph. This book owes its existence to an initial impulse provided by Christian Poitevin (Brussels) and Georges Balmino (Toulouse), and to the sponanteous enthusiasm of Heinz Draheim (Karlsruhe).

The publisher of my previous technical books, Dr. Christoph Müller–Wirth of Wichmann Verlag (Karlsruhe), eagerly took up the idea of this book and followed the different stages of its realization with his interest and advice. Several colleagues from various fields of science and philosophy kindly agreed to critically read the manuscript: L. Jonathan Cohen (Oxford), Franz Moser (Graz), Hellmuth Petsche (Wien), Wilfried Schröder (Bremen), Hans Georg Schwarzacher (Wien), Ivan Supek (Zagreb), and Hans–Jürgen Treder (Potsdam–Babelsberg). Their remarks helped remove some major errors and many minor mistakes and even misprints, besides providing additional information and inspiration. Their criticism and encouragement were absolutely essential. Of course I bear the full responsibility for the contents of the book and for any remaining errors.

Dr. Konrad Rautz (Graz) was of invaluable help in editing: he drew the figures, composed the index, helped in proofreading, and provided advice in difficult questions of word–processing. My secretary Mrs.

Ruth Hödl did all the word–processing in her usual efficient, dedicated, and painstaking way. Last but not least, my wife Gerlinde read various versions of the manuscript and was my adviser in questions of biology and theology, besides confirming that the book can be read also without mathematics. All this help is gratefully acknowledged.

Graz, February 1995

Helmut Moritz

Contents

Part A

Human Perception and Thinking

Chapter 1

The human brain

1.1 The brain and the nervous system

Philosophers, use your brain!
J.Z. Young

If there is an agreement about some fact in natural philosophy, it is about the fact that human thinking is inseparably related to our brain. It seems therefore appropriate to start lectures about natural philosophy with a review of some basic knowledge about the structure of our brain.

Very much is known in brain research, but in view of the huge complexity of the topic, it is still relatively little. Our present treatment further simplifies our knowledge to a few basic items which one should know in order to speak meaningfully about some topics in natural philosophy.

That the physiological or anatomical structure of our brain is at all relevant to philosophy, is not generally accepted in philosophy. The laws of logic and mathematics are so much more rigorous than our mainly empirical knowledge about brain structure, and these "laws of thinking" do not seem to depend very much on the architecture of our brain. Thus books on philosophy do not generally start with the physiology of the brain. It seems, however, that modern trends in the theory of knowledge, known by the name of evolutionary epistemology, do depend on mechanisms of perception (sight, hearing, etc.) and on the processing of the data of perception by the brain. The same holds for artificial intelligence and for other advanced aspects of automatic computation which, though not directly the subject of the book, do have some points of contact, such as algorithmic and axiomatic methods and their limitation by Gödel's theorem.

A brief introduction will be sufficient for our present purpose, however. Additional literature is abundant and may be consulted in case of interest. On the other hand, the reader should not be bothered by details which may confuse someone who is not used to think in terms of neuroscience. For a first reading it is suggested that the reader does not stop at some unusual details but skips them and continues to read in a cursory fashion.

This is generally recommended for reading the present book (and philosophical works in general). Don't get stuck at the first difficulty but continue reading. When you are through with the book, start again from the beginning: many difficulties will be clear now. If necessary, iterate the procedure, concentrating on the passages which you have found difficult. Never try to conquer a difficulty by force, but rather by patience.

The architecture of the brain

Main parts of the brain are

– the brainstem ("reptilian brain"),

– the limbic system ("mammalian brain") and

– the cerebrum ("conscious brain").

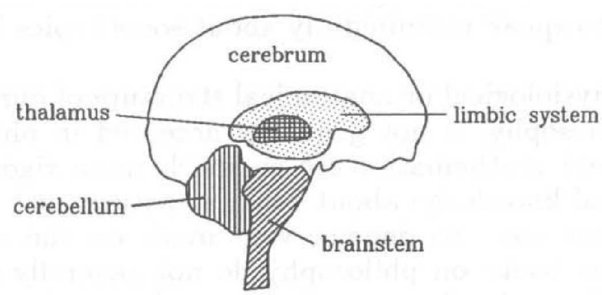


Figure 1.1: Main parts of the brain

The *brainstem*, from the point of view of biological evolution, is the oldest part of the brain (some 500 million years). With characteristic oversimplification, it is called *reptilian brain* because it looks somewhat like the entire brain of a reptile. It is mainly concerned with the most primitive functions of life support: control of breathing and heartbeat.

The *cerebellum* (Fig. 1.1) is an intermediate structure concerned with coordinating the activity of the muscles and maintaining bodily equilibrium (cf. sec. 4.1). It seems to be responsible for largely unconscious skills such as automobile driving or piano playing.

The *limbic system* is a ringlike structure which was the highest component in the brain of a reptile but is most highly developed in mammals. It regulates body temperature, blood pressure, etc. ("homeostasis"). It controls emotional reactions necessary for survival: appetite, agressiveness, flight reactions, and sexual feelings. It also seems to be basic for learning and memory (the so–called *hippocampus*).

The most important structure of the limbic system is the *hypothalamus* (located under the thalamus, Fig. 1.1). It is *the* regulator for body functions and states: eating, drinking, sleeping, waking, chemical balances, hormones, etc. It is the body's thermostat for blood temperature and acts by feedback, as also technological regulators do (see sec. 4.1).

The hypothalamus is intimately connected (by chemical and electrical messages) to the adjacent *pituitary gland*. This is the master gland of the body which regulates it by hormones, directly or by stimulating other glands to emit hormones.

The *thalamus* (Fig. 1.1) serves as a central relay station for external sensory, especially visual, information. This information is then relayed to certain areas of the cortex (see below) for final processing.

The *cerebrum* is the part of the brain which is specifically related to human intelligence. It consists of two hemispheres, the left and the right hemisphere, which are connected by the *corpus callosum*. There is a curious cross–connection: the *left* hemisphere receives information from the *right* half of the body (e.g., the right eye), and also controls the *right* half of the body (e.g., the right hand). Similarly, the right hemisphere is related to the left side of the body.

Generally speaking and oversimplifying, the left hemisphere mainly takes care of logical and analytic thinking and language, whereas the right hemisphere is chiefly responsible for intuitive, "synthetic", activities such as geometric intuition and music. Both hemispheres, however, closely cooperate and form "one system".

Most activities of the cerebrum go on at its surface, the *cortex*. The cortex is about 3 mm thick and is intricately wrinkled and folded, so that its huge surface fits into a relatively small skull.

Certain areas of the cortex are related to certain activities: the primary visual area, the auditory cortex, the body sensory areas, the motor cortex, and the speech areas (Broca's and Wernicke's areas). These areas, however, are by no means to be taken in a strict and exclusive sense; in a way, all activities affect the whole brain.

Corresponding to its function, the cortex is densely packed with nerve cells (neurons) whose axes (axons, see below) are *orthogonal to the surface of the cortex*, so that we have a structure reminiscent of level surfaces and plumb lines in geodesy. Rational thinking goes on mainly in the cortex; the limbic system (hypothalamus etc.) seems to contribute the emotional background.

Neurons

Each brain contains about 10^{11} nerve cells or *neurons*. Fig. 1.2 shows, very schematically, a typical nerve cell. The synaptic buttons (some 10^{14}!) attach to dendrites or somas of neighboring cells. There is a tiny "synaptic cleft" between a button and the next dendrite, which is bridged by chemicals (*neurotransmitters*) sent from button to dendrite.

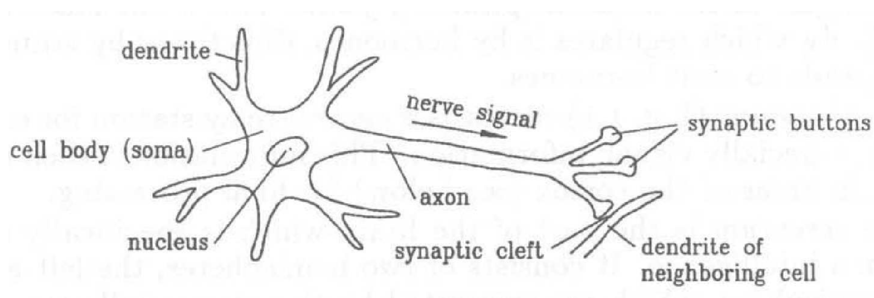


Figure 1.2: A typical neuron

The average length of neuronic fibers in the cortex is 4.1 km/mm^3!

How is the nerve signal transmitted along the axon? It by no means acts like a copper wire, say; rather, the transmission resembles a transversal seismic wave! A transversal wave oscillates in a direction *orthogonal* to the direction of propagation. Similarly, the nerve signal is a progressing impulse of diffusion of certain ions ($\mathbf{K}^+$, $\mathbf{Na}^+$, $\mathbf{Cl}^-$) *across* the tube–like membrane forming the axon. The rather complicated details of the process are without relevance to the present context.

The normal speed is only around 5 meters per second; the speed can be increased by a *myelin sheath* of the *axon* up to 100 m/s. Again, the reader is referred to the literature for details.

Important is the following fact: the "firing" of an impulse is *all or none*. The shape and size of the impulse is irrelevant; important is only whether the neuron fires or not. *Synapses* (consisting of synaptic buttons and synaptic clefts) may be *excitatory* or *inhibitory*. Their combined effect on the neuron determines whether it fires or not. The system: *all* (1) or *nothing* (0) is essentially digital, using a dual number system (consisting of zero's and one's) which is not unlike the dual system used by a digital computer. Very simple examples may be found in (Penrose 1989, pp. 393–394).

Each historical period compares the human body to its favorite machines: formerly it was a mechanical or heat engine, with a pump for the heart. Now it is fashionable to compare the brain to a huge parallel–processing digital computer. (There seems to be some "analogue computation" too, but on a minor scale, by "non–spiking neurons".) An important difference to present computers is a high redundancy: if part of the "brain computer" fails, other parts may take over.

Assumptions about the brain computation have lead to the construction of *artificial neural networks*, which, together with more conventional computing devices, play a considerable role in *artificial intelligence* (AI).

H. Petsche (Vienna) pointed out to me that the "digital firing" (all or none) mentioned above does not imply that the brain works exactly like a digital computer. From (Cohen and Stewart 1994, p. 454) we quote:

> Nerve cells don't "compute" digitally, but they do use discrete pulses to communicate over large distances.

That "brain computation" is not really "digital". For a detailed treatment of "neural computing" cf. (Kohonen 1988, Chapter 9: a very readable non–mathematical chapter in a otherwise rather technical book). Also the discussion in (Penrose 1989, pp. 392–399) is highly relevant.

Let us also not forget that "computation" or logical thinking is only a tiny fraction of our intellectual activity which also includes joys, desires, fears, poetry, music, etc., cf. end of sec. 2.1.

Suggested additional reading. The interesting reader may start with the didactic masterpiece (Ornstein and Thompson 1984) and continue with more technical books such as (Eccles 1973) or (Thompson 1985).

More advanced is (Oeser and Seitelberger 1988). The brain as a computer is treated very clearly in the introductory book (Churchland 1988) where also the elements of AI and artificial neural networks may be found. A profound study of AI (and many other things) can be found in the classics (Hofstadter 1979) and (Penrose 1989).

1.2 Brain and mind

The ghost in the machine.

Gilbert Ryle

The relation between brain and mind is one of the most controversial problems in philosophy. It is closely related to the problem of matter and mind.

At a first glance, there seems to be no problem at all. Our thoughts, feelings, and emotions are clearly *mental*, belonging to the mind. A tree, a house, or a stone are clearly *material*, consisting of matter: we can kick a stone, live in a house, or fall from a tree, breaking our leg. What could be more material?

Modern science gives a rather different picture of matter. It consists of extremely small molecules, separated by large distances. Molecules consist of atoms, and the atoms have a tiny nucleus with electrons orbiting around them like planets orbiting around the sun, and separated by similarly huge (in a relative sense) distances. So matter essentially consists of empty space. If my foot kicks a stone, an empty space kicks another empty space (to be honest, what matters are the forces that act in these "empty spaces").

If modern science makes matter less "material", it also makes mind less "mental", as the philosopher Bertrand Russell said. The study of our brain has shown us that, e.g., emotions are related to the activity of the (material!) limbic system, and even the most sublime manifestations of human thinking seem to correspond to the firing of certain neurons in the cortex.

A fair and didactically useful comparison between brain and mind seems to be possible by modern computer terminology: *the brain is the hardware, the mind is the software* of human thinking. The same software can be run on different computers, and the ideas contained in the present book (supposing that it contains any) can be understood by different readers, that is, by different brains.

This analogy: brain $\hat{=}$ hardware, mind $\hat{=}$ software, seems to be reasonably appropriate if we bear in mind that it is a crude picture only which should not be taken too seriously. Difficulties start when we ask for the *"ontological status"* of mind versus matter: is the mental software "essentially different" from the "brain hardware"? Some philosophers think that such questions make no sense, others think differently. Is the mind a "substance"? Dear reader, don't ask me what a substance is. Webster's New Collegiate Dictionary gives various definitions, e.g., "that which constitutes anything what it is". Do you understand this? I don't.

More seriously, a substance may be regarded as something that can exist independently of everything else, or at least independently of other "substances". In this sense, if mind, or "soul", is a substance, then it can exist independently of matter, and a consequence could be the *immortality of the soul*!

Dear reader, you will see how these questions are burdened with emotions. Far from being meaningless or purely theoretical phantoms, they have inspired noble actions of unselfish humanity as well as caused the most abominable wars.

But let us return to science. In a way, mind seems to *emerge* from the activity of the brain. On a lower level, life may be an emergent property of matter. "Emergent" means something new, qualitatively different. Here we might have an agreement.

If lifeless *matter* becomes sufficiently highly organized (containing complex organic molecules), then *life* may emerge, and if living brain tissues are particularly highly organized, the phenomenon of *mind* may emerge.

This seems to be acceptable to agnostics, dialectic materialists and Christians alike. Differences are again to be expected (and found) in ontological questions: in which way does mind, emerging from matter (the brain tissue), differ from matter? Is it only an "outgrowth" of matter or does it have an independent reality? ("Emergence" only implies that mind is "something new".)

Let us summarize the points which could be agreed upon by a majority of scientists, philosophers, and theologians:

(1) The mind is related to the brain somewhat as software is related to hardware; see above.

(2) Mind is a new aspect of brain activity that emerged when biological evolution led to the genus "man".

These formulations are deliberately vague so as to be widely acceptable (of course, we take biological evolution for granted).

So far, we have used the adjective "ontological" twice without defining *ontology*. Here Webster's definition is excellent:

"the science of being or reality; the branch of knowledge that investigates the nature, essential properties, and relations of being."

If you don't understand this, learn its use in the same way as a child learns to use new words: by (largely unconscious) practice. Just go ahead reading!

Theories of mind and brain

– *Materialism*: only material processes exist, mind is just the "subjective side" of brain processes but has no reality whatsoever. All "living" and "mental" processes can be reduced to changes in matter fully determined by the laws of physics (*reductionism*).

– *Epiphenomenalism*: what really goes on, are material brain processes; our subjective experiences are real but logically redundant (nothing is lost if we forget about mental phenomena).

– *Monism* (*or identity theory*): only *one* "substance" exists which, however, expresses itself in two ways: it has two "sides" or *aspects*: mind and matter. Mind and matter, so to speak, are the two sides of one and the same sheet of paper. This theory has been proposed by philosophers as different as Baruch Spinoza and Bertrand Russell. Closely related is

– *Panpsychism*: everything which is material (an atom or an electron, say), has also some mental or psychic aspect, however small. In living beings, and still more, in the human mind, these mental aspects are getting more and more coordinated. This theory seems quite strange at first glance; it is proposed, however, by some of the greatest philosophers, such as Leibniz and Alfred North Whitehead. It is, essentially, also implied in dialectic materialism, as well as (by definition) in the monism of Spinoza and Russell.

– *Idealism*: the opposite of materialism: everything that exists is an idea in a mind. The tree in front of me exists if I look at it: then it is in my mind. If I don't look at it, it does not exist, unless my wife happens to look at it. Since this is a rather queer world if trees would jump into and out of existence depending on someone's looking at them, George Berkeley (1685–1753) discovered a simple reason why the objects of this world continuously exist: because God is always looking at them. (The reader should not be misled by the author's occasional sarcasms which are by no means expressions of disrespect: Berkeley was a great philosopher!) This view is expressed in a limerick by Ronald Knox, quoted from (Russell 1945, article on "Berkeley"):

> There was a young man who said, "God
> Must find it exceedingly odd
>   If he finds that this tree
>   Continues to be
> When there's no one about in the Quad."
>
> REPLY
>
> Dear Sir:
>   Your astonishment's odd:
> *I* am always about in the Quad.
>   And that's why the tree
>   Will continue to be,
> since observed by
>   *Yours faithfully*,
>   God.

Plato, Kant, Fichte, and Hegel are also considered idealists, although in a somewhat different sense: they regard mind as a primary concept, and matter a concept derived from mind.

This view is not so far from opinions of some modern physicists who argue as follows. It is very difficult to define matter directly. For a physicist it is natural to say that matter is what satisfies the laws of physics. Now these laws are expressed by mathematical formulas, which are certainly mental rather than material structures. Similarly, the entities satisfying mathematical formulas are mathematical functions, which are equally mental. Hence matter is a mental construction.

Such an argument can be refuted, as almost all philosophical arguments can. Nevertheless it seems to incorporate at least a spark of truth. Finally, we shall discuss

– *Dualism*: matter and mind are essentially different concepts or, as some philosophers say, different substances. It is certainly the commonsense view. If I cut my finger (which is a material object), I feel pain (mentally). Here, clearly, *matter acts on mind*.

  This also works in the opposite direction. If I see an apple in front of me and wish to eat it (desire is mental), I take and eat the apple (a material object). Here, so to speak, *mind acts on matter*.

  Philosophers have wondered how "different substances", matter and mind, can act on each other. The first example, a cut in the finger causing pain, is so common–sense that no one uses to ask many questions about such a "direct causation". The opposite, a mental volition causing a bodily movement (taking the apple) has given, and is still giving, rise to innumerable discussions on "*downward causation*", mind acting on matter. We shall consider various aspects of this problem in sec. 4.1 and elsewhere in this book.

  Dualism in a normal, not extremist, view thus seems to be quite natural, and we shall use it in this way. It becomes a problem only if it is exaggerated to mean two "absolutely different" substances that are "absolutely separate" and hardly able to interact with each other. This view has been introduced into philosophy by René Descartes (Cartesius, 1596–1650) who is considered the founder of modern philosophy, besides having been a great mathematician (Cartesian coordinates!).

The danger with philosophy is that natural language is adapted to everyday use, but it has not been designed for philosophy. So philosophy must use every–day words, making them more precise. If this alleged precision is over–stretched and the new meaning of an old word is taken too literally, then we may overdo an argument and come to exaggerated conclusions. Alfred North Whitehead speaks of the "*fallacy of misplaced concreteness*". This may happen, for instance, if we consider mind as an indestructible substance, from which immortality necessarily follows. (I have no objection against immortality, on the contrary, but it is perhaps not quite so simple.)

Some followers of Descartes used the concept of God to mediate between the "noninteracting" substances, matter and mind. It seems that the concept of God is sometimes invoked to serve as a "*deus ex machina*" if a philosopher, overdoing his reasoning, has reached an impasse (see also the limerick above). It appears doubtful whether such a concept of God as a "Universal Metaphysical Problem Solver" who repairs fallacious reasoning of philosophers, is really adequate.

Thus it seems necessary to admit some interaction between body and mind (causation and downward causation as mentioned above). Thus body–mind dualism should actually be an *interactionism*.

This is the thesis impressively argued in the book (Popper and Eccles 1977), which is a fundamental (though not uncontroversial) reference on the mind–body problem. The authors also treat, extensively and rather fairly, the other alternatives mentioned above: materialism, epiphenomenalism etc. Another excellent review is provided by the papers by Sperry, Dewan and others in (Globus et al. 1976).

From p. 75 of (Popper and Eccles 1977) we take an argument against materialism concisely formulated by J.B.S. Haldane in 1932: "If materialism is true, it seems to me that we cannot know that it is true. If my opinions are the result of the [physical and] chemical processes going on in my brain, they are determined by the laws of [physics and] chemistry, not of logic." (Insertions between brackets are by H.M.) Remark: In order to be true, a proposition must be determined by the laws of logic.

This argument, which goes back to Epicurus (around 300 B.C.), is not entirely cogent, and it was later retracted by Haldane himself. (Determination by the laws of physics and determination by the laws of logic are not incompatible, as any electronic computer shows.) It is, however, a typical and elegant example of a philosophical argument. These arguments will be taken up again in sec. 6.4.

A more cogent argument against materialism is the fact that usually brain processes are considered on the basis of classical physics only (cf. Churchland 1988, otherwise an excellent introduction!). This, however, is totally inadequate since the microscopic definition of matter must be in terms of *quantum mechanics*. Here, however, matter and mind seem somehow to be intimately interrelated as we shall see in sec. 3.5.

At any rate, concepts such as materialism, idealism or dualism cannot be discussed in the narrow context of brain–mind interaction only. Therefore, we shall come back to these concepts later in this book, especially in sec. 5.1.

Metaphysics and ontology. At the present point, it might be useful to give a first, rather preliminary explanation of some very frequently used philosophical terms. *Metaphysics* studies philosophical questions which are beyond the reach of natural sciences, "beyond (Greek: *meta*) physics". (Some people regard as "metaphysical" even concepts or theories that *explain* rather than merely *describe* phenomena in physics itself.) Traditional metaphysical questions are the relation between God and world, the human soul, "what the world really is" and other "ontological" questions. *Ontology* (already touched upon above) studies the various forms of "being" or "existence", e.g. the nature and existence of matter, mind, mathematics, etc.

The validity of metaphysics, including ontology, is denied by *positivism*: only observed data and logical conclusions are meaningful. (These definitions are deliberately simplified: interpretations differ and there are many shades and nuances.)

Suggested additional readings. The problem of mind and brain in the context of quantum theory is discussed in several excellent books. If one takes (Penrose 1989) together with (Stapp 1993) and (Lockwood 1989), one should be pretty much up to date. It is recommended, however, to postpone reading these books until quantum phenomena have been treated in sec. 3.5. A more "classical" but highly regarded text is (Edelman 1989). An interesting update on (Popper and Eccles 1977) is (Eccles 1994), whereas (Hofstadter and Dennett 1981) provide a fascinating counterpoint. A nice anthology is (Rosenthal 1991). For ontology, there is the wonderful classic (Gilson 1972), but it is probably too difficult at this early stage.

1.3 Human perception

> *You don't see a pretty girl, you just see some colored patches.*
>
> Anonymous

Auditory perception

Although our main emphasis will be on visual perception, the sense of hearing also presents great interest. We listen to the voice of a friend, or to a piece of music.

The ear is an organ of extreme sensitivity. If it would be only a little more sensitive, we would constantly hear an unbearable background noise produced by the thermal motion of the air molecules. The inner ear contains a beautifully designed resonance mechanism which, so to speak, performs a very precise and detailed *harmonic analysis* of the auditory signal, a sound wave. The individual frequencies are heard distinctly and separately. If we strike a chord of four simultaneous notes on the piano, we hear all four frequencies distinctly in their true pitch (if we have an absolute ear) or in their correct relation of frequencies (which is enough for the average musician and music–lover). Thus if we strike on the piano simultaneously the notes c and e, we shall hear both of them distinctly and not "averaged" to an intermediate note d, say. This is important because in visual perception precisely this "averaging effect" occurs.

On the other hand, auditory perception in man does not give very precise information on the direction from which the sound comes, and on the location of the source, except when these faculties are especially trained in blind people. (Animals such as bats can "hear" directions very precisely!)

It is also worth noting that each half of the brain receives information from both ears.

Visual perception

This is the classical form of perception, on which philosophers traditionally have laid the greatest emphasis. It is well known that each eye acts like a small photographic camera. A lens generates a picture of the outer world on the *retina*, again usually a very precise picture, which is then transmitted by its optical nerve to the cortex. The main relay stations of both "optical tracks" are the two *lateral geniculate bodies* (LGB), one in the right and the other in the left part of the thalamus, cf. Fig. 1.1 on p. 4. The *left* hemisphere of the brain (or the left half–brain) receives its visual information from both eyes, but from the *right* half of the visual field of each eye! It is similar for the left visual field of each eye, from which information goes exclusively to the right LGB and the right visual cortex.

Optical processing occurs mainly in the retina, the LGB's and the visual cortex of both hemispheres.

Slight differences in the direction of the optical axes lead to small differences in the respective images on both retinas, called *parallaxes*, and they provide the third dimension, *depth*. This is called *stereoscopic vision*.

In view of the high state of present electronic *image processing*, it is tempting (but not cogent) to compare the processing of visual data of the eye with our image processing technologies, and stereoscopic vision has been employed already for several decades by *photogrammetry* for the same purposes, namely to construct a three–dimensional model from two–dimensional images.

Color perception. The eye's retina consists of closely packed receptors, some 10^7 cones and 10^8 rods. The *rods* are responsible for vision in conditions of faint illumination; they do not recognize colors. (Try to distinguish colors in a moon–lit landscape!) The *cones* are active in normal daylight. They consist of three types, one type having maximum sensitivity for a wave length of 430 nanometers (*blue*), the second type for 530 nm (*green*) and the third type for 560 nm (*red*).

We remember that visible light is electromagnetic radiation between wave lengths from about 400 to 700 nanometers (1 nm = 10^{-9} m).

If all wavelengths are uniformly represented, we speak of *white light*. If some wavelengths are prominent, we see a color corresponding to the average of these wave lengths. Thus the eye does not perform such a sophisticated and detailed harmonic analysis as the ear does with sound waves. That also could not be expected because then we should have a full harmonic analyzer (corresponding to one ear) at each of the 10^7 cones which is patently impossible. On the other hand, the retina provides precise two–dimensional and even, by the joint stereoscopic effect of both eyes, three–dimensional spatial position.

Any cone has maximum sensitivity for one particular wavelength and decreasing sensitivity for neighboring wavelengths. Every wavelength stimulates each of the three cone types to a certain extent, and the net impression is precisely the given wave length.

If the incident light contains several wavelengths, we thus see only *one "average" color*. It is as if simultaneously striking the two notes c and e on the piano we would hear only the intermediate note d! This fundamental difference between sight and hearing has already been mentioned at the beginning of this section; it means that there are no "color chords" in the sense in which there are musical chords (nevertheless, artists and critics speak of "harmonious colors"!). Thus music and painting are essentially different arts.

This averaging (or rather "mixing") of two wavelengths by the eye even produces a color to which no natural wave length corresponds: *purple* as the result of mixing blue and red. The natural spectrum reaches from blue over green, yellow and orange to red; it is a segment of a line. Purple "closes" this (physical) segment to become a (subjective) circle!

Thus, at least in color vision, there is an enormous simplification of wavelength information: a light ray that contains many, even infinitely many, wavelengths gives only *one* color. Color vision is a "many–to–one relation".

Oriented lines, moving objects, grandmother cells. Also in other respects, vision does not simply provide us with a one–to–one, so to speak, photographic image of the external world. There are cells in the cortex that react only on lines or stripes of a certain orientation: horizontal, vertical, or inclined at 45°. Other neurons react only if the object is moving. It is even said that there are single neurons that react only on a certain individual shape, such as the face of your grandmother ("grandmother cells").

Suggested additional reading: it is sufficient to mention two recent standard works written by outstanding neuroscientists in easily accessible language and treating different aspects: (Hubel 1988) and (Young 1987). We also mention the September 1992 issue of *Scientific American* dedicated to the present problem and in German: (Singer 1990, 1994).

1.4 The evolutionary theory of knowledge

Wär nicht das Auge sonnenhaft,
die Sonne könnt es nie erblicken.
Johann Wolfgang von Goethe

The "non–photographic" characteristics of vision mentioned at the end of the preceding section can be explained by the fact that the sense organs, as well as other parts of our body, have been shaped by biological evolution, passing through virtually all stages of the animal kingdom.

Hearing in animals was not designed for enjoying Beethoven's music. Hearing served to inform animals of approaching enemies or victims: briefly, *to help them survive*. With sight it was similar: vision in animals did not serve primarily to make them aware of the beauty of a landscape, but to help them recognize and distinguish *food and danger*. For this purpose it is essential if the object is moving or not, and also outlines and horizontal or vertical strips are important, as well as direct recognition of relevant objects: friends or foes ("grandmother cells", cf. end of sec. 1.3).

So the use of "sense data" for abstract purposes such as art, science, and philosophy came at a very late state of human development. Rather than blaming sense data for not providing us directly with all desired scientific information about the world, we should be surprised and grateful that sense data and their analysis by the human mind have taught us so much already and will (hopefully) continue to do so.

A priori and a posteriori knowledge

In the morning we go to our car, open it and turn the ignition key. Why? Because we know that, as a rule, the car will start and that we can then drive to our office.

The expectation we have *before* turning the ignition key is some kind of *a priori* knowledge (*a priori* means "beforehand"). If the car engine really starts *after* turning the key (which may not happen on a very cold winter morning), then we know "empirically" or *a posteriori* that our expectation has become true (*a posteriori* means "afterwards").

A priori information is generally believed to be provided by logic and mathematics: $3 + 2 = 5$, regardless whether we speak of apples or of human persons. It is also believed to be always true (although one cloud plus one cloud may coalesce to form another cloud: $1 + 1 = 1$ for clouds; this may be a joke, but not a trivial one!).

A posteriori information is obtained by empirical observation: watching the car start in our example.

As this example shows, a priori and a posteriori elements interact in almost every human activity, and clearly also in science: theories (*a priori*) are tested by experiment (*a posteriori*).

What was discussed in philosophy already before Kant but was placed in the center of philosophy by Immanuel Kant (1724–1804) was the respective role and certainty of the a priori and a posteriori elements.

Universally recognized by philosophers is only that logic is a priori and observation is a posteriori. Logically true propositions (tautologies) are also called "*analytic*", everything else (especially propositions about empirical facts) is called "*synthetic*".

Thus all analytic propositions are a priori, and all empirical propositions are a posteriori. Now Kant's question was "*Are synthetic propositions a priori possible*?" Take for instance the theorems of mathematics. They are independent of empirical observations, and were in fact considered *synthetic a priori* by Kant. Bertrand Russell (1872–1970) thought he could deduce all of mathematics from the axioms of pure logic and consequently believed mathematics to be a priori *but analytic*. Other contemporary mathematicians rather incline towards Kant's opinion.

Kant believed not only mathematics, but also our geometrical *and physical* three–dimensional space to be given *synthetically a priori*, as well as other "categories" such as time and causality.

He thought this to be true *absolutely*, not only approximately, *by the* *structure of our mind*. He believed to have performed a "Copernican revolution" in philosophy by putting the very basic laws of physics not into nature but into our mind: We simply cannot think of space differently than as three–dimensional and Euclidean.

Since Einstein's General Theory of Relativity (or even since the non–Euclidean geometries of Gauss, Bolyai and Lobačevsky) we know that Kant was not right in this respect. Nevertheless the discussion of *a priori* and *a posteriori* continues to the present day. See also our outline of Kant's philosophy in sec. 5.4.

Eddington's example. Eddington (1939, p. 16) gave a nice illustration of a scientific *a priori*. A marine biologist is exploring the life of the ocean. In order to get specimens of animals living in the sea, he throws a net and examines the catch. He discovers

(1) No sea animal is less than 5 cm long.

(2) All sea animals have gills.

In order to make sure his discovery is correct, he repeats this experiment several times at various places. His laws are always confirmed, so he concludes that they are universally true.

It is obvious that at least Law (1) is not an objective law of nature, but a consequence of the experimental setup. It would have been different had he used a net of a smaller (or larger) mesh size.

This illustrates that we can hardly avoid subjective, "a priori", elements. Sense perceptions thus are as much *given* ("sense data") as *fabricated* by our brain's highly complex processing (Eddington's net!).

Evolutionary epistemology

Epistemology is nothing else than "*theory of knowledge*". The famous biologist Konrad Lorenz (1903–1989) was the last professor on the philosophical chair of Immanuel Kant in Königsberg. Studying Kant's philosophy, he recognized that Kant's *a priori* conditions for human knowledge could be identified with man's perceptual structures (eye, ear, etc.) developed in the course of human evolution. Man's world view is largely, though by no means exclusively, conditioned by his physical and mental constitution. For instance if, instead of perceiving, in the electromagnetic spectrum, only the wave lengths from 400 to 700 nm, we would be able to visualize also other frequencies, then our visual representation would be quite different. (We do perceive infrared, but as heat.) Or think of the other extreme, a blind person.

This theory beautifully explains why our perceptions agree so well with nature (e.g., sight, hearing, and touch give consistent results). The old philosophical question of "*adaequatio mentis ad rem*" (correspondence of mind to nature) is thus simply explained: if an animal's sense impressions do not correspond to reality (e.g., if it hunts after a nonexisting prey), it will not survive since it does not conform to Darwin's principle of "survival of the fittest". It and its potential successors will have long ago been eliminated from the process of evolution.

The animal's *a priori* is its perceptual apparatus fit for survival, so to speak, the set of its "working hypotheses for survival".

In his important book "*Filozofija znanosti i humanizam*" (Philosophy of Science and Humanism, SNL, Zagreb, Croatia, 1991, p. 85), Ivan Supek shows that the evolutionary theory of knowledge goes back to the famous Austrian physicist Ludwig Boltzmann (1844–1906).

Science. Science is man's extension of his perceptual apparatus; primarily also for survival, later, since the times of the ancient Greeks, with the emergence of "intellectual curiosity", for satisfying the desire to know for its own sake. Theories are no longer (only) "working hypotheses for survival", but "working hypotheses for understanding nature". The primary *a priori* is, of course, constituted by logic and mathematics, on one hand, and by *the structure* of our sense perception, on the other hand. In which way logic and brain structure are related, is still almost entirely controversial. Our perceptual apparatus, on the other hand, has certainly been provided to us by evolution.

A secondary *a priori* for scientific research are the theories of physics, chemistry, biology, etc. In contrast to Kant, however, they are no longer regarded as absolute truths, but as working hypotheses of varying degree of certainty, and always subject to "falsification" or, at least, to revision; cf. sec. 6.5.

The bucket and the searchlight. Modern philosophers have emphasized to active role of the subject in information gathering. The logical positivists around Bertrand Russell considered material objects as "logical constructions from sense data", these data being more or less passively connected in a "bucket". (The expression is Popper's (1979, Appendix 1).) This view certainly suffers from Whitehead's "fallacy of misplaced concreteness", oversimplification. It is simply not true that we perceive "a red patch surrounded by green", no, we immediately recognize it as a rose. (It needs a high–level philosophical abstraction to see a red patch where everyone else sees a rose.)

The modern theory regards the perceptual apparatus as a "searchlight" to discover what we presume, expect, hope for, or fear, such as a hunter watching for a lion. As we have mentioned at the beginning of the present section, this is also the reason why movement detectors in the visual retina are so important.

But scientific theories and hypotheses, too, serve as such searchlights for exploring nature. We hope to *verify* them by experiment, or if we follow the severe ways of Karl Popper, we must try to *falsify* them, like a merciless professor who tries to fail all students except the very best ones (sec. 3.9).

Let us briefly recapitulate our main points. Logic and mathematics seem to be absolutely *a priori*. The scientific theories are also used in an *a priori* way, as working hypotheses, but they are subject to refutation or revision by experience. (See also Quine's model at the end of sec. 4.5.)

The logical–positivist theory of sense data received more or less passively, from which the external objects are obtained by "logical construction", does not correspond to reality. The *a priori* elements, from primitive fears and expections to the most advanced scientific theories, work as searchlights for discovering and understanding our external world.

Already animals perceive their foes or preys *directly as external objects* (in however rudimentary a way). They certainly do not construct their enemies mentally from sense data: they would be dead long before having completed the logical construction. Also man perceives directly, not sense data, but external objects. This is at least what man believes, and the "*hypothetic realism*" of "evolutionary epistemology" asserts he is basically right.

If I remember correctly, Alfred North Whitehead said that we would not even notice an elephant unless we would expect to meet this animal. This is another example of the searchlight thesis. To give a more everyday example: if one hikes through nature with a botanist or an ornithologist, one is surprised how much can be seen with their help which otherwise would pass unnoticed.

Additional reading. As an elementary and excitingly written introduction we recommend (Ditfurth 1976), as well as (Young 1987). On a medium level is the classic (Lorenz 1973). Philosophical books on evolutionary epistemology are (Popper 1979) and (Vollmer 1990). The subject is treated from somewhat different angles by Maturana and Varela (1987) and Piaget (1970). The best book, combining readability and great depth, on all problems of philosophical epistemology, including a fine chapter on Kant, is (Hartmann 1965).

The view of logical positivism is beautifully described in the booklet (Russell 1929). We recommend it because a person interested in philosophy must also be able to understand and appreciate a different point of view.

We shall come back to the philosophical aspects of this problem repeatedly later in this book. So at this point, only Ditfurth and Young can be understood fully. We also point out once more that the present book should, in principle, be understandable without additional reading, which is of course recommended for those desiring to know more about certain problems.

Chapter 2

Logic and mathematics

2.1 Elements of symbolic logic

You are not thinking:
you are just being logical.
Niels Bohr

Set theory

The notion of *set* should nowadays been known from school. (In the present treatment, we shall pursue an intuitive approach, like in school, sacrificing full logical rigor to simplicity.) A set consists of *elements*. It can be *discrete* or *continuous* (Fig. 2.1).

Figure 2.1: A discrete set (A) and a continuous set (B) in the plane

A set is equivalent to a *property*. Consider the poet Shakespeare. Some of his properties are
(M) He was a person (a human being).
(A) He lived in England.
(B) He was a poet.
(C) He was born in 1564.

Instead of *property* (M) we may also say: he is an element of the *set* of all persons; instead of (A): he belongs to the set of all beings who lived in England; instead of (B): he is an element of the set of all poets, etc.

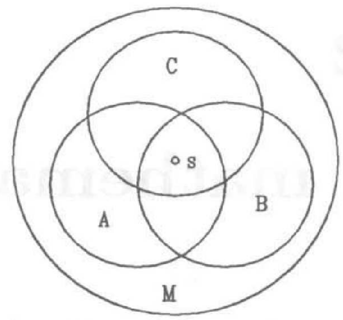


Figure 2.2: Shakespeare (s) belongs to the intersection of the sets M (persons), A (English), B (poets), and C (persons born in 1564). (Discrete sets are conveniently represented by continuous sets in the plane!)

This may be symbolically represented as follows (Fig. 2.2). Shakespeare, denoted by s, is an *element* of the set M, symbolically

$$s \in M \quad . \tag{2.1}$$

Obviously also $s \in A$, $s \in B$, and $s \in C$. The set A is a *subset* of the set M, symbolically

$$A \subset M \quad . \tag{2.2}$$

Obviously also $B \subset M$ and $C \subset M$. (The sign "$\in$" holds for elements and "$\subset$" for subsets!)

We also need the concepts of *union*

$$A \cup B \tag{2.3}$$

and *intersection*

$$A \cap B \tag{2.4}$$

(Fig. 2.3).

From Fig. 2.2 we see that

$$s \in A \cap B \quad , \quad s \in B \cap C \quad , \quad s \in A \cap C \quad ,$$

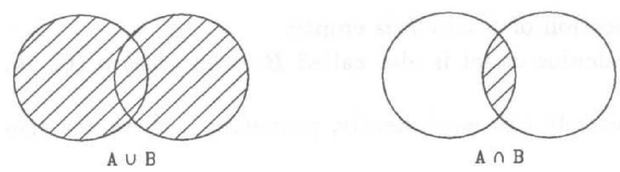


Figure 2.3: Union and intersection

$$s \in A \cap B \cap C \quad , \quad s \in M \quad , \quad s \in M \cap A \cap B \cap C \quad ,$$
$$A \cup B \subset M \quad , \quad A \cap B \subset M \quad , \quad A \cup M = M \quad ,$$
$$A \cap M = A \quad ,$$

etc.

We also need the concept of complementary set (with respect to some comprehensive set M), see Fig. 2.4. In an obvious notation we may write

$$\overline{A} = M - A \quad . \tag{2.5}$$

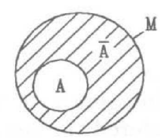


Figure 2.4: The complement of a set

If M denotes the set of persons (human beings), we may also form the complement $\overline{M}$ with respect to some more comprehensive set N, say the set of animals: $\overline{M} = N - M$. We may read $\overline{M}$ as "not M". For instance

$$n \in M$$

if n denotes the Roman emperor Nero (however inhumanly he behaved), but

$$n \in \overline{M} \quad \text{or} \quad n \notin M$$

(read: n is not an element of M) if n denotes Nero as the name of a dog.

We also need the concept of *empty set* $\emptyset$, e.g. the set of all triangular squares. Thus if T is the set of all triangles and S the set of all squares, we have

$$S \cap T = \emptyset \quad ,$$

the intersection of S and T is empty.

The calculus of set is also called *Boolean algebra* (G. Boole, 1815–1864).

Besides sets (or, equivalently, properties), there are also *relations*, e.g.

$$aRb \quad . \tag{2.6}$$

The two individuals a and b are connected by the relation R, e.g.

$$\begin{aligned} &a = 2, \quad b = 3, \quad R = < \text{ (smaller than):} \quad 2 < 3; \\ &a = \text{son}, \quad b = \text{father}, \quad R = \text{younger than} \ . \end{aligned}$$

In the latter case "aRb" means "a son is younger than (his) father".

There are many excellent introductions to set theory, e.g. (Halmos 1960).

Numbers

Let us try to define the integers 1, 2, 3, ... Consider the number 2. Obviously, "2" may be considered the property which is common to all pairs:

$$2 = \text{set of all pairs} \quad .$$

This set–theoretic definition (maintained, e.g., by the famous logician and philosopher Bertrand Russell) appears counterintuitive, because the set of *all* pairs is clearly infinite, but remember that "set" has been seen to be equivalent to "property", and "common property of all pairs" is intuitively acceptable. But does not the concept of "pair" presuppose the number "2"? The definition seems to be *circular* (*circulus vitiosus*, vicious circle).

We may, however, proceed "recursively" as follows. Let S_1 be a set of one element only, or to avoid using "one", a set of consisting of any individual. For individuals a, b, c, ..., S_1 consists only of a or only of b or only of c, ... S_1 is called a *one–set* here (the name "unit set" is used more frequently).

Then

$$\begin{aligned}
1 &= \text{set of all one–sets;} \\
2 &= \text{set of all two–sets, i.e., of all sets which} \\
&\quad\ \text{become one–sets by removing one element;} \\
3 &= \text{set of all three–sets, i.e., of all sets which} \\
&\quad\ \text{become two–sets by removing one element;} \\
&\quad\ \text{etc.}
\end{aligned} \tag{2.7}$$

Those readers who do not like the expression "set of all ...", may replace it by "property of all ...". (For a more rigorous version of the argument cf. (Carnap 1958, p. 71).)

Thus we can recursively define all positive integers. *Negative integers* and zero can be defined as pairs of positive integers:

$$-2 = 1 - 3 = (1,\ 3) = (5,\ 7) = \cdots$$

since $-2 = 5 - 7$ etc. Similarly,

$$0 = 1 - 1 = (1,\ 1) = (2,\ 2) = \cdots$$

Rational numbers are also defined as pairs of integers:

$$\frac{3}{7} = [3,\ 7] = [6,\ 14] = [9,\ 21] \quad \left(\frac{9}{21} = \frac{3}{7}\,!\right) \text{ etc.}$$

Irrational numbers are defined as infinite sets of rational numbers, e.g.,

$$\pi = \{3,\ 3.1,\ 3.14,\ 3.141,\ 3.1415,\ 3.14159\ldots\} \quad ; \tag{2.8}$$

this is a set of rational numbers, e.g.,

$$3.14 = \frac{314}{100} \quad ,$$

which better and better approximate π. Thus arithmetic is reduced to logic (set theory). Through analytical geometry in Cartesian coordinates (see also sec. 2.6), geometry can be expressed in terms of arithmetic; the coordinates are positive or negative, rational or irrational numbers.

In this way mathematics can be reduced to logic. This is at least what G. Frege and B. Russell thought (around 1900). The standard work is "*Principia mathematica*" by B. Russell and A.N. Whitehead

(1910–1913); it is, however, far too difficult to be included in the list of suggested additional reading. We recommend (Carnap 1958).

This reductionist procedure is not without problems because the concept of *infinity* enters; cf. (Weyl 1949) and (Barrow 1992).

Logic of propositions

A *proposition*, denoted by p, q, r, …, is a sentence or a statement. E.g.

$$
\begin{array}{lll}
p & \ldots & \text{It rains.} \\
q & \ldots & \text{Today is Friday.} \\
r & \ldots & \text{The street is wet.}
\end{array}
$$

Propositions can be *true* or *false*. In the first case, we assign the truth value T or simply 1, in the second case, F or simply 0.

The following symbols are used:

$$
\begin{array}{rl}
p \wedge q & p \text{ and } q \quad (\text{both } p \text{ and } q) \\
p \vee q & p \text{ or } q \quad (\text{either } p \text{ or } q \text{ or both}) \\
\sim p & \text{not } p \\
p \Rightarrow q & p \text{ implies } q \quad (\text{if } p, \ \text{then } q) \\
p \Leftrightarrow q & p \text{ is equivalent to } q \quad (p \text{ holds if, and only if, } q)
\end{array}
\tag{2.9}
$$

There are so–called *truth tables*, e.g.

$$
\begin{array}{c|c}
p & \sim p \\ \hline
1 & 0 \\
0 & 1
\end{array}
\qquad
\begin{array}{c|cc|l}
\vee & 1 & 0 & \leftarrow q \\ \hline
1 & 1 & 1 & \\
0 & 1 & 0 & \\ \cline{1-3}
\uparrow & & & \nwarrow \\
p & & & p \vee q
\end{array}
\qquad
\begin{array}{c|cc|}
\wedge & 1 & 0 \\ \hline
1 & 1 & 0 \\
0 & 0 & 0 \\ \hline
\end{array}
\qquad
\begin{array}{c|cc|}
\Leftrightarrow & 1 & 0 \\ \hline
1 & 1 & 0 \\
0 & 0 & 1 \\ \hline
\end{array}
\tag{2.10}
$$

It is clear that, if p is true, then $\sim p$ is false, and vice versa. This explains the first truth table.

To understand the second table remember that, in our example, $p \vee q$ denotes

"It rains or today is Friday."

This is true if p has the truth value 1 (regardless of the truth value for q). This gives the first line of the second truth table. If p has the truth value 0 (false), then $p \vee q$ is true (1) if q is true (1) but false (0) if q is false (0). This gives the second line of the table.

The reader might wish to check this by means of the concrete example just given.

The third truth table can be explained in a similar way; note that it is formally identical to the multiplication table for the numbers 1 and 0. This was one of the starting points for Boole's logical algebra!

The fourth truth table is almost self–evident: $p \Leftrightarrow q$ is true if $p = 1$, $q = 1$ or $p = 0$, $q = 0$ and false otherwise. ($p = 1$ is an abbreviation for "the sentence p has the truth value 1".) Other truth tables can easily be established.

We have the basic laws (which should be checked by means of our special example):

$p \vee \sim p$	There holds always p or not–p; p is either true or false (nothing else): *law of the excluded middle*, "*tertium non datur*"	
$\sim (p \wedge \sim p)$	p and not–p cannot hold simultaneously: *law of contradiction*	
$\sim (\sim p) \Leftrightarrow p$	double negation = affirmation	(2.11)
$p \Rightarrow p \vee q$	"If it rains, then it rains or today is Friday"	
$p \wedge (p \Rightarrow q) \Rightarrow q$	"*modus ponens*" (we can put — *ponere* — q as true)	
$(p \Rightarrow q) \wedge \sim q \Rightarrow \sim p$	"*modus tollens*" (we can remove — *tollere* — p as false)	

By means of these and similar laws, all logical conclusions (deductions, proofs) can be performed in a purely formal manner (e.g., also by an automatic computer). Therefore we speak of *formal logic*, in our case of *propositional calculus*.

We thus have a "mathematization" of logic, in a similar way as set theory has provided a "logization" of mathematics. Formal logic is understood to consist primarily of set theory and the calculus of propositions.

There is a close connection between these two branches of formal (or *symbolic*, or *mathematical*) logic, expressed in a correspondence between $\wedge$, $\vee$ in propositional logic and $\cap$, $\cup$ in set theory:

$$\begin{aligned} x \in A \cap B &\Leftrightarrow x \in A \wedge x \in B \quad , \\ x \in A \cup B &\Leftrightarrow x \in A \vee x \in B \quad . \end{aligned} \tag{2.12}$$

Here $x \in A$ (x belongs to the set A) is regarded as a proposition, say p, and $x \in B$ is another proposition, say q.

A final note on terminology: logical truths such as (2.11) are called *analytical truths* or *tautologies*. They hold always, independently of the special meaning of p or q.

Logical atomism versus holism

Symbolic logic works well if we deal with a finite number of well–defined and well–separated objects, such as apples in a basket. Thus it is no accident that Bertrand Russell (1929, p. 12) represents a "*logical atomism*" according to which the world consists of such discrete objects. Clearly, 1 apple + 1 apple = 2 apples, so that ordinary arithmetic holds.

But much in the world is continuous, with gradual transition, without sharp boundaries (cf. sec. 2.4). Consider, in a blue sky, two clouds which gradually coalesce into one cloud. Here we may say: 1 cloud + 1 cloud = 1 cloud, so that $1 + 1 = 1$ here! Generally, we may consider everything on earth related to (or correlated with) everything else; this is called "*holism*": the world is one interconnected whole; *the whole is more than the sum of its parts.* (The Greek word "*holos*" means "whole".) Following meteorology, it may well be that a butterfly flying over a meadow in Austria may cause a tornado in Florida ... This is Edward Lorenz' "butterfly effect" well known from chaos theory (sec. 3.2).

If holism were absolutely and exclusively true, then we could not speak of anything without dragging in the whole universe. (If we think of Mr. Smith, we must also consider the house in which he lives, generally his surroundings, his city, his country, the planet Earth, the solar system, ...!) So the truth seems to be somewhere in the middle between logical atomism and holism.

Logic and linguistics

Some modern philosophers have considered the object of logic, and the task of philosophy in general, to analyze ordinary language ("linguistic analysis"). They identify human thinking with (internal) speaking. This identification seems to be questionable, however: mathematicians "think" in terms of indistinct images and structures (Hadamard

1945; Penrose 1989, p. 423), and what about people who are equally fluent in several languages? In which language do they think? My answer, based on introspection, is: in none. I also think generally in images and indistinct structures (unless I prepare explicit formulations in a certain language for lectures or for writing).

It seems, however, that linguistic formulations play, in philosophy or history, a much larger role than in mathematical sciences. This may be one of the reasons why English is used as a *lingua franca* almost universally in natural sciences but much less so, e.g., in history. Philosophy is often difficult to translate, and poetry is frequently almost untranslatable.

We do no wish to downplay the importance of language for human thinking. It has been claimed that a child who has not learned to speak *one* (any) human language by the age of about 12 years, will forever be unable to think as a normal adult (Cohen and Stewart 1994, p. 355). Language is indeed an indispensable prerequisite for human thinking.

Thus there is no doubt that linguistic analysis is important in logic and philosophy, but it seems to be a tool rather than a goal in itself.

Logic and intellectual activity

Reading some philosophical books, one could get the impression that human intellectual activity is "pure thinking" expressed in objective language, logic being "the science of correct thinking".

A moment of reflection is sufficient to see that this is by no means the case. Saying "Of course, you are right!" may express almost everything, from objective agreement to furious sarcastic disagreement, depending on the way we are saying it. A poem expresses much more than any grammatical analysis can show. *emotions* pervade almost everything we can say or think. Our *limbic system* (sec. 1.1) is constantly active indeed!

Much of our intellectual activity is nonconscious or subconscious. Assume we give a speech without using a manuscript. During the moments when we express an argument, our subconscious keeps "thinking ahead" on what we are going to say next. At the same time, we are trying to suppress our anger about a letter which we just received, helped by the (largely unconscious) expectation of a good lunch after the lecture ... So many lines of "thinking" are going on simultaneously: our mental activity is indeed a "highly parallel processing", to use computer language which is now, more than ever, seen to be totally inadequate for describing the activity of our brains ...

Our expression need not be verbal. Getting a red face may express our embarrassment better than any words; and furiously walking out of the room could be our final argument. The behaviorist psychologists may not be so wrong after all!

Earlier we have talked about nonverbal thinking in mathematics. What about the performance of a Beethoven sonata by a pianist? In sec. 1.1 we have seen that the cerebellum is heavily involved. Does this mean that the performing pianist is not active intellectually? If you think so, don't tell it to her; otherwise you may get a painful nonverbal answer!

These examples from elementary psychology show that "verbal thinking" is indeed only a small fraction of our total intellectual activity. They are not meant to discredit logic but to place logical thinking into its proper, lofty but highly abstract, place.

2.2 The axiomatic method

What is truth?

Pontius Pilate

Axioms are basic propositions from which all true statements of a certain branch of science of mathematics can be derived by a purely formal procedure (also by an automatic computer!).

The first and best–known axiom system is Euclid's axiom system for elementary geometry (around 300 B.C.). A complete and rigorous axiom system for this purpose was given, however, only by David Hilbert in 1899.

Geometrical statements are fully proved only if they are derived in a purely formal way from the axioms, without using intuition or figures. Figures, etc., are to be considered only as "heuristic" aids in guessing mathematical theorems or making them plausible; a rigorous deduction from the axioms must then follow. A purely formal way of deduction is frequently called an *algorithm*.

An axiom system may satisfy the following requirements:

(1) *Consistency*: this is the absence of internal contradictions; it is absolutely necessary.

(2) *Completeness*: all true statements that can be formulated in our logical system, can be derived from the axioms. This requirement is often satisfied but there are also incomplete axiom systems.

(3) *Independence*: the axioms should not be redundant. This requirement is desirable but not absolutely necessary.

A complete axiom system gives, so to speak, "implicit definitions" of the concepts which it contains. For instance, "points", "straight lines", or "planes" are those mathematical entities which satisfy the axioms of geometry (they are by no means identical with small dots or "lines" made with chalk on the blackboard!).

Examples for incomplete axiom systems are Peano's axiom set for integers or the axioms of group theory (this remark is only for the specialist).

The axiomatic method accounts for the great abstractness of modern mathematics (e.g. the famous French school of "Bourbaki"). Axiomatization is also a goal for other "exact" sciences such as physical theories. Many axiom systems can be found in (Carnap 1958).

However, axiomatization is a final goal, but never the beginning of a science. In physics, but also in parts of mathematics such as differential geometry, theorems and whole theories are first derived intuitively or heuristically, making use of figures, additional assumptions, etc. The same holds for differential and integral calculus, where the basic theorems were first derived intuitively, in an inexact way (using "infinitesimal small quantities"). Only at a later stage they were made rigorous by limit processes. The most advanced theories of physics (string theories, supersymmetry, Feynman integrals) are still largely at a heuristic stage.

Even the "simplest" mathematical discipline, *arithmetic* or *number theory* (the theory of the properties of natural numbers 1, 2, 3, ...) *cannot be fully based on a single axiom system*. This is Gödel's theorem to be treated in the next section.

2.3 Logical paradoxes and Gödel's theorem

Even one of their own prophets has said:
"Cretans are always liars."

St. Paul's Epistle to Titus

Logical antinomies

Already around 1900, Russell, analyzing work by Gottlob Frege (1848–1925), met with the first difficulties in reducing mathematics to logic (more precisely, to set theory). He found *antinomies* or *paradoxes*, that is, logical contradictions.

Russell's antinomy: Let M be the set of all sets which do not contain themselves as elements. Does M contain itself? Answer: it contains itself as element, if and only if it does not contain itself! Thus the concept of the set M is obviously contradictory.

This requires some abstract thinking. A more concrete form has been given to it also by Russell: A barber in a village shaves exactly those of its male inhabitants who do not shave themselves. Does the barber shave himself? Yes, exactly then if he does not shave himself!

Russell's antinomies and similar paradoxes show that the concept of "set" in its "naive" form may be contradictory. Various methods have been tried in order to avoid such antinomies:

(1) Limitation to simple logical systems in which it is impossible to even *formulate* such antinomies.

(2) *Axiomatic set theory*: the axioms should exclude "dangerous" sets. However, such attempts so far have been largely "ad hoc". Even after having excluded known antinomies, one cannot be sure that later on one cannot find new antinomies in some remote corner of the system. So far, the consistency of axiomatic set theory has not been proved!

Hilbert's program of "*formalism*" (around 1920) was intended to give formal calculi for logic and mathematics whose consistency was to be proved according to mechanical rules (as realized, e.g., by an automatic computer).

Gödel's theorem

Hilbert's program received a death blow in 1931 when the Austrian logician Kurt Gödel proved that such a program is impossible.

In particular, Gödel proved the following: In a certain formal ("computerizable") system which comprises logic and arithmetics, a proposition G (or logical formula) can be constructed, which *asserts its own unprovability*. Is G provable? If so, then G is true and therefore, according to what it asserts, unprovable! Thus both G and $\sim G$ are provable within the system.

Now, if both a formula (G, in our case) and its negation ($\sim G$) can be derived, the underlying axiom system is *inconsistent* (i.e., logically self–contradictory). If arithmetic is consistent, then neither G nor $\sim G$ can be derived from the axioms.

Now comes the big surprise: by "meta–mathematical" reasoning *outside* the logical system based on our axioms, *it can be shown that G is true*! We, so to speak, manage to get at a level *above* the formal system, and "looking down" we recognize that G must be true. (This, of course, is not so easy as it sounds!)

What have we found? If arithmetic is consistent, then G is *true* and hence *unprovable* (remember that G asserts its own unprovability!). There is a true statement G in the system which cannot be derived from the axioms! According to the definition of completeness given in sec. 2.2, *arithmetic, if consistent, is incomplete*!

Now the consistency of arithmetic has been proved by other means ("transfinite induction", Gerhard Gentzen, around 1936). Hence *arithmetic is incomplete*; it cannot be fully described by a (finite) system of axioms! This reveals a fundamental limitation of the power of the axiomatic method.

Because of the essential incompleteness of arithmetic, it has been considered possible that some well–known unsolved problems such as Goldbach's conjecture ("Every even number can be represented as the sum of two odd primes") or even Fermat's last theorem ("The formula $x^n + y^n = z^n$ has no solution for integer x, y, z, n except for $n = 2$") may be a consequence of the fact that the presently known axioms for arithmetics are incomplete and hence insufficient to derive them.

An important consequence of Gödel's and related theorems is that important areas of mathematics (and hence physics and other sciences involving mathematics) cannot be completely characterized by axiom systems, they are no "calculi" or "*algorithms*" computable by machines. *Human reason transcends computability in an essential way.*

Gödel's paper is extremely difficult. Fortunately, there exists a wonderfully lucid presentation (Nagel and Newman 1958).

The antinomy of the liar. Gödel's proof reminds one of the "antinomy of the liar" which goes back to ancient Greece:

"What I am now saying is false"

(or simpler, "I am lying now"). Is this sentence true or false? You have guessed correctly: it is true if it is false and it is false if it is true (Fig. 2.5).

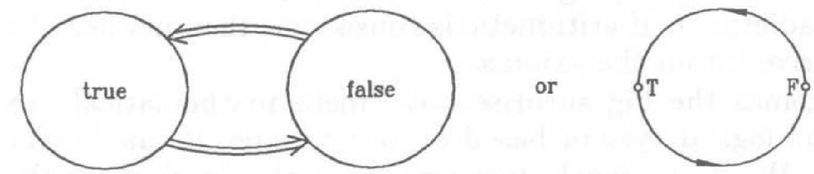


Figure 2.5: The "vicious circle" in the paradox of the liar.

This paradox, attributed to the Cretan Epimenides, was known to Aristotle and Cicero, and is even alluded to St. Paul's letter to Titus (1,12–13). It is said that some ancient "philosopher" lost his sanity by meditating on this sentence, which becomes plausible on looking at the "vicious circle" of Fig. 2.5 which could really drive one crazy.

Gödel replaced "false" by "unprovable" and hence avoided Epimenides' trap by a hair's breadth. He converted the explosive energy contained in Epimenides' paradox into a powerful argument which by logicians is considered the most important single achievement of 20th century logic.

Self–reference

The common feature of all these paradoxes is self–reference: Russell's set contains itself, the liar asserts something about what he himself is just saying, etc.

It seems that self–reference has an importance far beyond some esoteric paradoxes: its applications reach from human thinking ("the thinking thinks the thinking") and artificial intelligence to cosmology. Hofstadter (1979, 1985) has written two remarkable heavy volumes about it. Self–reference also seems to lie at the basis of dialectic thinking. Thus we shall meet it frequently in our book.

Let us here mention the delightful little book (Smullyan 1980), which is a veritable treasure of paradoxes.

The logical antinomies and Gödel's theorem seem to indicate that even formal logic and mathematics cannot be made arbitrarily precise. We may speak of a *Gödelian uncertainty* of logic and mathematics as an analog to the well–known *Heisenberg uncertainty relation* in physics (sec. 3.5).

Foundations of mathematics

We have already met two ways of founding arithmetic or the theory of *natural numbers* (positive integers), which may then serve as a basis of other branches of mathematics:

(1) *Logicism*: arithmetic can be reduced to logic (set theory), cf. (2.7) in sec. 2.1.

(2) *Formalism*: logic and mathematics can be constructed together by "mechanical rules" (Hilbert's program), see above in the present section.

There is a third approach:

(3) *Intuitionism*, which goes back to the Dutch mathematician Dirk Brouwer (since 1907). It by no means implies that formal mathematical proofs are to be replaced by intuitive reasoning, quite on the contrary: The logical requirements of intuitionistic proofs are so stringent that a large part of contemporary mathematics cannot be proved by intuitionistic methods and would have to be sacrificed.

Gödel's proof affects logicism and formalism. It does not affect intuitionism since Gödel's argument could not even be formulated within intuitionism.

This does not mean that intuitionism is complete, it is even much more incomplete than "classical arithmetic" (logicist or formalist approaches)! Therefore it is a historic curiosity rather than a method applied by contemporary mathematics (as embodied in the Bourbaki school).

The main reason for all the problems and paradoxes in logic and mathematics is *infinity*. Russell's set is "highly infinite"; self–reference implied "zero distance": the barber shaves himself (he has "zero distance from himself"), the liar says something about which he is just saying (both statements are coincident or have "zero distance"), and "zero", or "infinitely small", is also an aspect of infinity, at least in the present context (remember that one frequently bluntly puts $1/0 = \infty$, although mathematicians may frown on it).

The problem why consistency proofs in set theory are so difficult is that "strongly infinite" sets occur. Brouwer's intuitionism tries to avoid this problem by admitting only *potential* infinities, such as the series 1, 2, 3, ... which potentially goes to infinity, but forbidding *actual* infinities such as the set of *all* infinitely many integers.

At any rate, many complicated systems of logic and set theory have been constructed in order to avoid known paradoxes, but we can never be sure that, some day, we may not run into a new antinomy.

It is true that few working mathematicians care much about these "esoteric" difficulties: usually "nothing happens". The great mathematician Weyl (1949, p. 235), however, writes about the arbitrariness in present logico–mathematical systems: "How much more convincing and closer to facts are the heuristic arguments and the subsequent systematic constructions in Einstein's general relativity, or the Heisenberg–Schrödinger quantum mechanics."

Modern references include (Rucker 1982) and (Barrow 1992).

2.4 Inexact concepts, "fuzzy logic"

Fuzzy logic makes better washing machines.
Advertisement

Inexact concepts

Exact are certainly the concepts of mathematics: numbers such as 1, 2, 3, $\sqrt{2}$, π; geometrical points, straight lines, circles, spheres, etc.

Rather exact are empirically defined physical concepts such as "precisely" drawn lines, circles, "physical points" realized, e.g., by the intersection of two empirical lines (+) or as the center of small circles (∘) found on any illustration in a book on geometry. Rather exact are also the concepts of "logical atomism" (see the end of sec. 2.1): Mr. Smith, his house, a lamp or a book on his table, Mr. Smith's dog, an apple, and a huge amount of similar "well–defined" concrete objects.

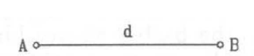


Figure 2.6: Distance between two points

"Rather exact" are also physical, astronomical, or geodetic measurements. Consider a distance d between two "points" A and B (Fig. 2.6).

Clearly, even the "points" A and B are not defined physically with absolute precision: on paper, they are defined not much better than to an accuracy of 0.1 mm; in nature, perhaps by a cross (+) on a stone to an accuracy of 1 mm; most precisely, by marks engraved on metal or glass to an accuracy of about 0.001 mm. At any rate, the distance d cannot be measured to an accuracy higher than the accuracy to which physical points are defined. Let the result of our measurement be 23.281 m. If the points are defined to 1 mm only, the result may equally well be 23.280 or 23.282 m. In geodesy, we may write 23.281 m $\pm$ 1 mm, there the "standard error" is denoted by $\pm$ 1 mm. This is not precisely equivalent to the statement that the measurement lies in a "confidence interval", say between 23.280 and 23.282 m, but we may disregard this distinction for the present purpose; cf. also sec. 4.4.

Here we already see that classical principles of logic such as the *law of contradiction* or the *law of the excluded middle* (sec. 2.1, formulas (2.11)) no longer hold with empirical concepts: $p \vee \sim p$, "d is either 23.281 or not", is violated, there may as well be $d = 23.281$ m or $d = 23.2812$ m ($\neq 23.281$ m). The same holds with the law of contradiction: $\sim (p \wedge \sim p)$: $d = 23.281$ m (p) and $d = 23.2812$ m (not–p) may hold simultaneously.

Inexact are many concepts of ordinary life, starting from our body: nose (where in our face is the boundary line between our nose and an adjacent cheek?), finger (where on our hand does it begin?), but also a tall building (how high?), a hiking trail in the mountains (which to his embarrassment, the hiker frequently finds very difficult to follow), smoke, a cloud (see also the end of sec. 2.1), or the Earth's atmosphere in general (at its "edge" there are less and less molecules in a cubic meter; if there are only a few stray molecules or ions, is this still atmosphere or already empty space?).

Fuzzy sets

Many sets in ordinary life are not very precisely defined, e.g., the Earth's atmosphere just mentioned, considered as a set of molecules. Such sets with ill–defined boundaries are called "fuzzy" (Fig. 2.7).

Obviously, a cloud is a fuzzy set of water droplets. But also more abstract sets in ordinary life are "fuzzy". Think of the set of honest people. Who belongs to it? Saints, yes; thieves, no; but what about myself, whose behavior frequently is less than perfect? Or consider the set of red apples: which shades of red are permitted? Is an apple still red or already closer to orange? What about a red apple with some little yellow spots?

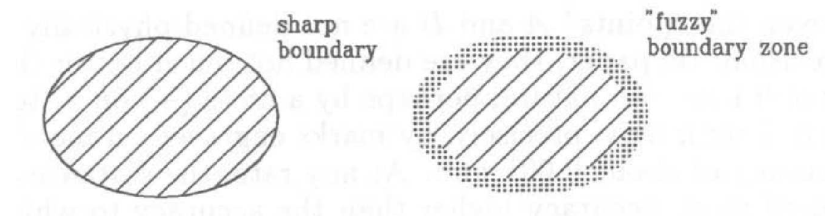


Figure 2.7: Precisely defined and fuzzy sets

Since, as we have seen in sec. 2.1, sets are equivalent to properties ("honest", "red"), also properties of everyday life are frequently fuzzy: think of "tall", "cold", "fast" and very many others.

Fuzzy propositions

The same "fuzziness" also applies to propositions (statements) of everyday life. A few examples: "It is hot today" (which temperature?), "You are stupid", "Mahler's symphonies are great".

Among the most popular unprecise statements are weather forecasts. "There will be light rain showers tomorrow." In the U.S.A., newspapers are frequently more explicit: "There is a probability of 25 % for rain tomorrow". Intuitively this is pretty clear, but what does it mean precisely? Will rain fall in 25 % of the next day (consisting of 12 h, say), that is, will we have $25\ \% \times 12 = (25/100) \times 12 = 3$ hours of rain tomorrow? If instead it rains for 4 hours, will we cease to trust our newspaper or the weather forecast in general?

Obviously not, we have to live with unprecise statements and try to interpret them reasonably, guided by experience. Much more difficult is the task to "precisely" formulate these "unprecise" statements, especially if we want to simulate human thinking by a computer. Frequently it does not work to consider some facts as precisely true and contradictory facts simply as false, as we ourselves unfortunately do often enough, following the example of small children ("*I* am right!" "No, *I* am right!")

Fuzzy logic

There are various cases and possibilities of formalizing these "fuzzy" concepts. Probably the oldest one (Gauss and Legendre, around 1800) is

(1) *Least–squares adjustment* used in geodesy, astronomy, physics, and other disciplines. This may be illustrated by a simple example. In a triangle we measure the three angles

$$\begin{array}{rcl} \alpha & = & 75^\circ 13' 37'' \\ \beta & = & 49^\circ 26' 13'' \\ \gamma & = & 55^\circ 20' 16'' \\ \hline \text{sum} & = & 180^\circ 00' 06'' \end{array} \tag{2.13}$$

Now we know that the sum of the angles in a plane must be 180°. Thus there is a contradiction.

The simplest way is to leave α and β and to change γ to $55^\circ 20' 10''$. This would correspond to the frequent human attitude, mentioned above, to considering some facts as precise and disregarding others. However, it contradicts our sense of fairness and justice: why not treat all observations in the same way, "adjusting" them equally? Thus the "misclosure" of $6''$ above 180° is divided evenly among all angles, diminishing each by $2''$:

$$\begin{array}{rcl} \alpha & = & 75^\circ 13' 35'' \\ \beta & = & 49^\circ 26' 11'' \\ \gamma & = & 55^\circ 20' 14'' \end{array} \tag{2.14}$$

so that now $\alpha + \beta + \gamma = 180^\circ$. A general procedure for such problems is *adjustment by least–squares estimation* . Being quite an elaborate method, it is beyond the scope of the present book (see, however, sec. 2.6). What is important is that we have a well–defined procedure which can be implemented on a computer.

(2) *Subjective probability* . According to classical logic (sec. 2.1), propositions are either true (truth value 1) or false (truth value 0). Now many statements are only more or less probable, as the example of a weather forecast has shown.

More about the concept of probability will be said in sec. 3.3, but even here it appears plausible to "interpolate" between the truth values 0 (impossibility) and 1 (certainty), defining as *probability* P of a sentence any number

$$0 \leq P \leq 1 \quad . \tag{2.15}$$

It is, of course, often difficult to assign, to any proposition under consideration, a precisely defined number P, in much the same way as it is often difficult to assign a definite grade G (in Austria: $1 \leq G \leq 5$, with "1" denoting the best grade and "5" meaning "failed") to a student in an examination, especially if it is required to give non–integer grades, such as 1.3. Still, decimal places are useful if average grades etc. are computed. In the same way, probabilities of rain of 0.2 (or 20%) or 0.9 (90%) are more telling than "some rain may be expected" or "rain is very probable".

If the probability of certain basic propositions (axioms) are known or assumed, then the probabilities of derived propositions can be computed (cf. Carnap 1950; Jeffreys 1961).

The choice of initial probabilities is often not so crucial; what is essential, is that "*subjective logic*" (so called because it refers to propositions which are considered subjective, rather than to objective facts) can be handled by a fixed algorithm. Thus, in much the same way as least–squares adjustment, it can be implemented on a computer.

(3) *Theory of fuzzy sets*. This theory, somewhat similar but not identical to probability theory, was developed since 1965 by L.A. Zadeh and others, mainly in the U.S.A. The practical breakthrough, however, was made in Japan, in the early '80s, where it was applied in a technologically innovative way. "By 1985 Hitachi had installed the technology's most celebrated showpiece, a subway system in Sendai, about 200 miles north of Tokyo, that is operated by a fuzzy computer" (TIME, September 25, 1989). It gives an astonishingly smooth ride and uses less energy than conventional systems.

Since then, "fuzzy control methods" are used in Japan and elsewhere to improve auto–focus cameras, washing machines, fuzzy shower systems that avoid too cold and too hot temperatures, etc.

As an example, let us mention a hypothetic automatic car braking system based on the following very "fuzzy" rules:

1. IF the *road curvature* is LOW
   AND
   IF the *speed* is LOW to NORMAL
   THEN the *braking* should be SOFT.

2. IF the *road curvature* is NORMAL
   AND
   IF the *speed* is HIGH,
   THEN the *braking* should be NORMAL.

3. IF the *road curvature* is HIGH,
   AND

IF the *speed* is VERY HIGH,
THEN the *braking* should be STRONG.

It is indeed astonishing that efficient and reliable control mechanisms can be developed on the basis of such vague rules.

It is perhaps not surprising that in an Eastern country like Japan, in which thinking is more inclined to "holism" (see end of sec. 2.1) should have exhibited such a spontaneous interest in "fuzzy thinking", whereas the West, in which "logical atomism" comes more natural, has at first been rather reluctant.

All types of methods based on inexact knowledge are comprehensively treated in the book (Spies 1993), from which we also have taken the example of "fuzzy automobile braking". The book (Kosko 1993) reads particularly well.

It seems that also the neural network of our brain, with its huge number of intricately connected neurons, may be well equipped to handle fuzzy information. Is precise thinking related to predominantly digital functioning of our brains, and fuzzy thinking to analog computation also going on in the brain (cf. sec. 1.1)? There may well be some relation of this kind, but it is almost certainly not that simple. See also end of sec. 5.2.

Informal reasoning

The arguments we hear in daily life, in quarrels and disputes, even in university lectures, are hardly capable of being expressed in the symbolism of formal logic as, for instance, mathematical proofs may be.

Curiously enough, the same holds also for the "informal reasoning" in philosophical arguments. Even philosophical concepts such as causality or determinism, matter and mind, freedom and law, are by no means sharply defined. They subtly change their meaning during a discussion. Sometimes this is intentional, sometimes it passes unnoticed. Lucas (1970, p. 58) speaks of "*chameleon words*".

In philosophical and other discussions, statements are not in general simply true or false. There are arguments *pro* and *con*, some carry great weight, some arguments are rather weak. Discussions may be intended to prove the opponent wrong, but this seldom happens. Mostly the participants in discussions and the readers of philosophical books are invited to follow the arguments, to appreciate their strength, validity, and cogency, and finally to form their own opinion.

Nobody could imagine to replace, in a discussion, the two opposing philosophers by opposing computers. So far, no philosophical book has been written (authored, I mean) by a computer. (Of course, philosophers use text processing, but this is a completely different matter . . .)

The informal reasoning which we admire in good philosophical books, is, by no means, "purely logical" as an innocent reader might think. As Niels Bohr has said (motto of sec. 2.1), thinking is much more than "just being logical". Only read a page of Sir Karl Popper or J.R. Lucas and analyze the structure of their arguments. They differ from dull logical deduction as a piano sonata by Beethoven differs from a monotonous finger exercise.

2.5 Dialectic thinking

Not–being is a form of being.

Plato

Introduction

We may define *dialectics* as that aspect of human thinking that transcends "algorithmic reasoning" which could as well (or better) be performed by a computing machine. A beautiful example has been pointed out by J.N. Findley in his article "The contemporary relevance of Hegel" in (Findley 1963). It is *Gödel's theorem*.

Formalized thinking based on axioms and, in principle, performable by a machine, is called an *object language*, or *symbolic language* or *formal language*. The informal language in which we speak *about* the operations of formal logic, is called a *metalanguage* (considerations *about* mathematics are called metamathematics!). In the present book, the metalanguage is (Austrian) English. It is a useful convention to consider the metalanguage to be of *higher level* than the object language.

Using this terminology, we may describe Gödel's proof (sec. 2.3) as follows.

In a certain symbolic object language, a proposition G is constructed which asserts its own *formal* unprovability. By *informal* reasoning, however, it can be shown that G nevertheless is *true*. Let us quote Findley (1963):

> But the unproveable sentence at the same time soars out of this logico–mathematical tangle [of the formulation of G in the object language] since the proof of its unproveability in *one* language [the object language] is itself a proof of the same sentence in *another* language of higher level [the metalanguage], a situation than which it is not possible to imagine anything more Hegelian.

(Italics are Findley's, explanatory insertions between brackets [ ] are mine.)

Never mind if you did not get all of the argument at first reading, just go ahead! This example, trying to show that highly respectable contemporary logical concepts are behind dialectics, was mainly intended to counteract the preconceptions against dialectics due to its dogmatic over–use (and sometimes misuse) in some contexts.

Already Plato identified dialectics with philosophy as the highest of all sciences, even above mathematics.

Its explicit use in classical German philosophy started with I. Kant (1724–1804), reached a first culmination with J.G. Fichte (1762–1814), especially in his Wissenschaftslehre of 1804, continued with F.W.J. von Schelling (1775–1854), and reached its final culmination with G.W.F. Hegel (1770–1831). For details cf. (Hartmann 1960), (Gulyga 1990) or (Kuznecov 1981), a very readable introduction is the booklet (Müller 1974).

Dialectics is also basic in Eastern philosophy, cf. (Capra 1975) and (Smullyan 1977).

Dialectics is primarily a logical discipline and has been so understood by all philosophers from Plato to Hegel. Karl Marx and Friedrich Engels tried to apply it to "nature" itself. This is not necessarily wrong: also modern physics applies logic and mathematics to nature. We shall, however, consider here dialectics as a logical discipline, independent of our metaphysical background (idealism, materialism, dualism, etc.; cf. sec. 5.1).

After this lengthy introduction let us try to describe dialectics first in the usual simplified manner. It is the triadic scheme

$$\text{thesis} \longrightarrow \text{antithesis} \longrightarrow \text{synthesis} \tag{2.16}$$

or better

$$\begin{array}{ccc} \text{thesis} & \longrightarrow & \text{antithesis} \\ \searrow & & \swarrow \\ & \text{synthesis} & \end{array}$$

The synthesis (S) is not simply a compromise between thesis (T) and antithesis (A), but S is a higher–level standpoint from which both T and A become understandable and even compatible. Thus, in the triad, the synthesis S, so to speak, lies on a higher level than both T and A. (*It would be absolutely wrong to regard A as the simple logical opposite* $\sim T$ (non–T) *in the straightforward sense of formal logic*, cf. sec. 2.1.

Dialectics is not that simple! Contradictions and opposing views do play a role, as we shall see, but not of the simple type p versus $\sim p$.)

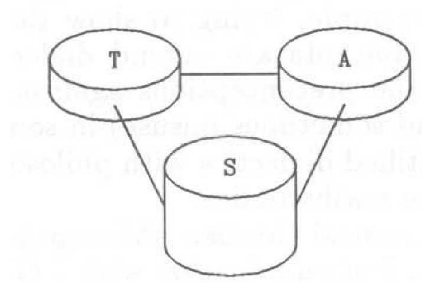


Figure 2.8: S lies on a higher level than T and A

It is important to remark right away that the usual scheme thesis — antithesis — synthesis is well suited for introductory and didactic purposes, but *it constitutes by no means the essence of the dialectic method.* Much more important is *rising to a higher level*, cf. Gödel's argument mentioned at the beginning of this section and Fig. 2.10 on p. 52. In short, *dialectic thinking is metathinking.*

Let us illustrate this by examples from science.

Example 1.

Light as wave (T) — light as particles (photons) (A) — quantum theory incorporating both aspects (S); cf. sec. 3.5.

Example 2.

The universe has always existed (Aristotle) (T) — universe has started at a certain time (Bible, "big bang") (A) — both propositions unified by a mathematical transformation ($t \to \log t$, Milne, sec. 3.7) (S)

Example 3.

Geocentric world system (T) — heliocentric system (Copernicus) (A) — equivalence by general relativity (Einstein) (S). This equivalence is a result of Einstein's *principle of general covariance*, according to which all reference systems are theoretically equivalent; cf. sec. 3.4.

Example 4.

Logical atomism (T) — holism (A) — actual world probably unifies both aspects (S); cf. secs. 2.1 and 2.4.

Example 5.

A geodetic example: measurement of all three angles α, β, γ in a triangle (sec. 2.4).
Thesis T: α, β, γ have been measured.
Antithesis A: $\alpha + \beta + \gamma \neq 180°$ (contradiction!).
Synthesis S: least–squares adjustment as described in sec. 2.4.

As we have seen above (Fig. 2.8) the important property is that *the synthesis lies on a higher level*. This is also illustrated by an example from everyday life.

Example 6.

A discussion in which two participants defend two apparently contradictory positions (T and A). A skilful chairman manages to convince both partners that their opinions, if seen from a proper perspective, are really compatible (S). If all three participants have good will, this usually succeeds, often surprisingly, giving all a profound feeling of satisfaction, especially if the initial positions T and A have appeared quite contradictory.

Let us return to empirical science. All measurements are affected by measuring errors, and no physical theory can be expected to hold absolutely. We shall now give an example for a measurement and one for a theory (cf. secs. 4.4 and 3.4).

Example 7.

T: A measured distance d has the value 18.85 m.
A: This cannot be true because of measuring errors, rounding–off, etc.
S: $d = 18.85\text{m} \pm 0.007\text{m}$. This shows that it is an empirical value and at the same time estimates its accuracy.

Example 8.

T: Newtonian mechanics holds exactly.
A: No, since relativity theory gives better results.
S: Newtonian mechanics holds to a certain accuracy which is defined by relativity theory ($v/c \ll 1$, cf. sec. 3.4).

The dialected process can sometimes be iterated in order to come closer and closer to reality, as the following two examples show.

Example 9.

A trivial example from mathematics. What is the value of $\sqrt{2}$?

$$T: \qquad \sqrt{2} = 1.4 \qquad\qquad A: \qquad \sqrt{2} = 1.5$$

Both answers are wrong. A better value is

$$S: \qquad \sqrt{2} = 1.41 \quad .$$

If this is not precise enough, we may iterate the procedure, taking (S) as a new thesis T_1.

$$\begin{array}{lll} S = T_1: & \sqrt{2} = 1.41 & A_1: \ \sqrt{2} = 1.42 \\ S_1 = T_2: & \sqrt{2} = 1.414 & A_2: \ \sqrt{2} = 1.415 \\ \quad S_2: & \sqrt{2} = 1.4142 & \\ \quad \text{etc.} & & \end{array}$$

Note that the synthesis is not the "cheap" solution $\sqrt{2} = 1.4145$, the arithmetic mean between T_2 and A_2. There is no "automatic" way of finding the synthesis. This always requires creativity or superior knowledge.

Such an iteration may be illustrated by Fig. 2.9, the right–hand figure showing convergence particularly well.

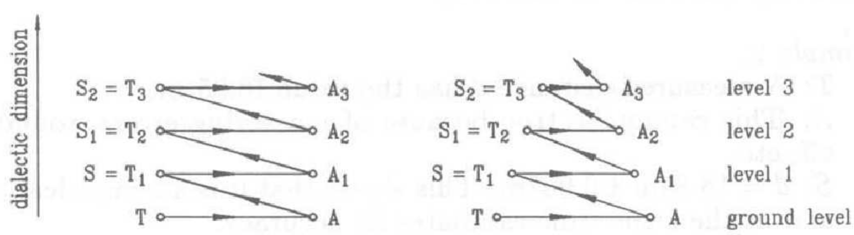


Figure 2.9: Illustrating different levels

Example 10.

Iterative solution of an equation $x - \phi(x) = 0$, assuming $\phi(x)$ to change only slowly with x.

T: A certain approximate value x_0 gives a solution.
A: $x_0 - \phi(x_0) \neq 0$ (contradiction!)

$$\begin{array}{lll} S = T_1: & x_1 = \phi(x_0) & \text{is a better solution} \\ A_1: & x_1 - \phi(x_1) \neq 0 & \text{(contradiction)} \\ S_1 = T_2: & x_2 = \phi(x_1) & \text{is again better} \\ A_2: & x_2 - \phi(x_2) \neq 0 & \\ S_2: & x_3 = \phi(x_2) & \\ \text{etc.} & & \end{array} \tag{2.17}$$

This is again illustrated by Fig. 2.9, especially by that on the right–hand side.

These primitive examples (except perhaps Example 6) have only served to illustrate the machinery of dialectics. Were it only for them, however, we should not need dialectics! There must be deeper reasons.

Characteristic features of dialectics

(1) *Opposites and contradictions* are an essential feature in nature and human life. They cannot be removed; on the contrary, they serve as driving forces. Examples:

freedom — law
accident — necessity
content — form (e.g., in art)
analysis — synthesis
self–preservation — care for others
conservation — progress
subject — object
matter — mind
justice — love
continuous — discrete
theory — practice
simplicity — complexity
competition — cooperation
order — chaos
and many others.

Human thinking is full of contradictions. A well–known example from psychology: you get a high reward if in the next ten minutes you will not think of a blue elephant. You would never in your life think of a blue elephant, but when you desire not to do so, you will certainly think of a blue elephant!

Following the physicist and Nobel Prize Winner Niels Bohr (1885–1962), such pairs of natural opposites are called *complementary*. The term "complementarity" comes from quantum mechanics, cf. Example 1 above and sec. 3.5. Bohr extended the *principle of complementarity* to other phenomena in nature and in the human spirit and has contributed essentially to making dialectic thinking respectable to scientists. Examples of complementarities in biological and other systems can be found in (Haldane 1939) and (Holzmüller 1984, p. 118).

(2) *Human language* (including logic and mathematics) is too crude and too inflexible to represent complex situations by a single linguistic perspective. Various different perspectives must be used.

Example 11.
"This leaf is green." Is it really green?
T: "The leaf is green."
A: "The leaf is not green."
S: "The leaf is green with small yellow and brown dots."

In a discussion it is frequently possible to bluntly assert the contrary of what the predecessor has said (even if it is not a "profound truth" in the sense of Bohr, see below). This may help render the discussion more precise; in many cases it can cause effective blank surprise. (Try it if you are sure of yourself!)

Human language is usually adequate for simple "discrete" objects such as apples. Elementary language, and even more so symbolic logic, however, may run into difficulties already with "simple" natural properties such as "green". The "green" of the leaf is not the ideal "green" of formal logic. Therefore even in Example 11, A is not $\sim T$ in the formal sense since "green" in T and A have slightly different meanings. Remember what we have said on "fuzzy" concepts and "chameleon words" in sec. 2.4.

Language is even less adequate for the abstract concepts of philosophy. This at least partly explains the multitude of philosophical systems which may be considered attempts for looking at the world from various angles, cf. sec. 6.8. Language and its inadequacy is a central topic of Wittgenstein's later philosophy; cf. sec. 5.4.

(3) *Rising to a higher level*. Another primary characteristic of dialectic thinking is "*thinking about thinking*", higher–level thinking about original thinking, above called "metathinking". This is done when discussing, in plain English (metalanguage), a computer program performing a logical calculus (object language), or even in the informal explanations in a book about symbolic logic such as (Carnap 1958).

> Hegel's dialectic corresponds to the sort of informal, non–formalizable passages of comment and discussion in a book like *Principia Mathematica* [by Russell and Whitehead], rather than its systematic text, and it has the immense importance of that interstitial comment.

(Findley 1963, "The contemporary relevance of Hegel").

This "metathinking" is also of essential importance in the formation of the human personality. Who has never learned to question his own actions and his own thinking, has missed the great opportunity to become a human person. Formal "thinking" can also be performed by machines and higher animals; only self–conscious "thinking about thinking" is probably restricted to man only. We again remind the reader what we have said about the dialectic nature of Gödel's proof at the beginning of the present section.

"Thinking about thinking" may involve several or even many levels in the way of Fig. 2.9. A striking example, due to Poul Martin Møller, may be found in (Bohr 1963, p. 13). A Danish student says:

> I get to think about my own thoughts of the situation in which I find myself. I even think that I think of it, and divide myself into an infinite retrogressive sequence of "I"s who consider each other. I do not know at which "I" to stop as the actual, and in the moment I stop at one, there is indeed again an "I" which stops at it. I become confused and feel dizziness as if I were looking down into a bottomless abyss, and my ponderings result finally in a terrible headache.

According to Rosenfeld, this is a "delightfully humorous illustration of Hegelian dialectics" (Folse 1985, p. 54). It also strongly reminds of Fichte (Versuch einer neuen Darstellung der Wissenschaftslehre 1797, II(2); Werke Band I, p. 526, simplified and modernized):

> When you say that you are conscious of yourself, you distinguish your thinking "I" from the "I" about which you are thinking: the "I" as a *subject* (S_1) from the "I" as an *object* (O_1). But in this process you are necessarily regarding the subject S_1 as the *object* O_2 of a new *subject* "I" (S_2), and so on: $S_1 = O_2$, $S_2 = O_3$, $S_3 = O_4$, etc. *ad infinitum.*

(4) *The geometry of dialectics*. If we represent syntactic "thinking about objects" by a movement in a plane, then the dialectic triad (T, A, S) certainly does not correspond to a triangle lying *in* this plane. The synthesis S rises *above* it (Fig. 2.8). To use the same metaphor, dialectics corresponds to a third dimension, rising above the basic plane of formal logic. This procedure may even be iterated, cf. Fig. 2.9.

The triad, or triangular motion, is by no means a necessary characteristic of dialectic thinking, as we have already mentioned. Instead of the triad ($T - A - S$) Fichte frequently uses a 5–term pentad ($T - A - TA - AT - S$), where TA is a partial synthesis with the accent on T, and AT similarly with the accent on A, and the mathematician Speiser (1952) uses a 7–term process in his attempt at a modern reconstruction of Hegel's "*Logic*". Weizsäcker (1992) uses the term "Kreisgang" (circular movement), in which the observer, so to speak, walks around the object in a circle, to look at it from all perspectives, at the same time rising higher and higher, like in a circular staircase. Geometrically, this corresponds to a spiral line or *helix* (Fig. 2.10). The projection of this helix onto the basic plane is, of course, a circle (Fig. 2.11), on which the projection of S coincides with T although, of course, S is on a higher level than T (Fig. 2.10).

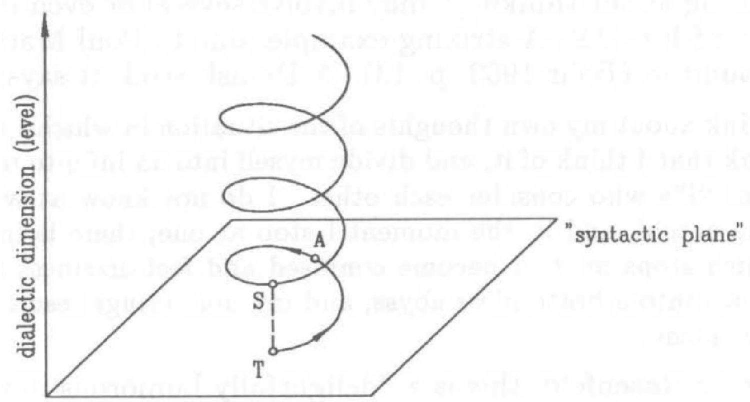


Figure 2.10: The dialectic helix

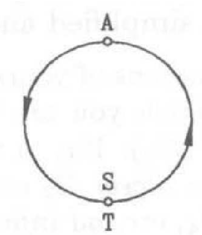


Figure 2.11: The dialectic helix only *appears* to be a vicious circle

Thus the dialectic helix is not simply a "vicious circle" corresponding to the paradox of the liar, cf. Fig. 2.5 on p. 36, in which we remain in the same plane, so that S really coincides with T. The liar's paradox and Gödel's proof correspond to each other very much as a circle and a helix (or perhaps the part of the helix in Fig. 2.10 which goes from T to S).

The dialectic helix represents one way to get out of a vicious circle: by rising higher and higher. Another way in which a vicious circle can be made to lose its viciousness is *iteration*: here going around continuously decreases the radius; the circle becomes a spiral (Fig. 2.12), which

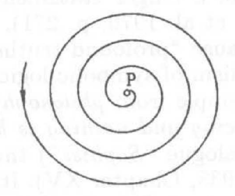


Figure 2.12: An iterative spiral

contracts to the desired solution P. If we regard the iterative solution of the equation $x - \phi(x) = 0$ (Example 10 above) as a dialectic process, then we go higher and higher with each iteration (Fig. 2.9, right), but this is not really essential because we avoid the vicious circle anyway by decreasing the radius to zero. So the iteration procedure described in Example 10 can be considered to remain on the same mathematical level.

A completely different geometrical interpretation of dialectic logic is found at the end of sec. 2.6.

Fundamental principles of dialectics

(I) *Unity of contraries*. Synthesis is the unification of thesis and antithesis at a higher level, or more generally, the unification of conflicting tendencies at a higher level. This principle has been known at least to Plato, who uses it in his dialogue "*Parmenides*" (Speiser 1959). At the end of the middle ages, Cardinal Nicolaus Cusanus (1401–1464) formulates the principle of "*coincidentia oppositorum*" (coincidence of opposites). Long after Fichte and Hegel, the physicist and philosopher Niels Bohr, when awarded a title of nobility for his scientific merits, chose the inscription "*contraria sunt complementa*" (contraries are complementary) as motto for his coat of arms. As a matter of fact, we have met Niels Bohr above as the pioneer of complementarity.

Dialectic thinking came natural to Niels Bohr: "The opposite of a trivial truth is falsehood. The opposite of a profound truth may well be another profound truth." Or even more concisely: "A deep truth is a truth whose opposite is also a deep truth." Both statements are condensed versions of a longer statement in (Bohr 1958, p. 66); the first follows (Globus et al. 1976, p. 271). (Again A is not simply $\sim T$ of formal logic, because "profound truths" are too complex to be expressible by the formalism of symbolic logic.)

The most famous example from *philosophy* is Hegel's fundamental triad: *the synthesis of being and nothing is becoming*. It is based on

Plato's saying (in the dialogue "*Sophist*") that "not–being" is a form of "being" (Whitehead 1933, Chapter XV). It is also a central topic of Eastern Philosophy (e.g. the I Ching). Other examples of "profound truths" in the sense of Bohr will be given later in this book.

We have already mentioned the case of *psychology*. The human personality is a synthesis of many conflicting and contradictory trends. The personality is the stronger, the more it is capable of synthesizing contradictory tendencies, conflicting desires, and contrasting experiences into a harmonious whole, thereby rising to a higher level of experience: "Was uns nicht umbringt, macht uns stärker". (What does not destroy us, makes us stronger.)

(II) *Negation of negation*. As we have pointed out at several occasions, the antithesis A is not the simple operation $\sim T$ (not–T) of formal logic. Hence the negation of the negation does not simply reproduce the thesis T, but gives a synthesis S on a higher level as we have remarked repeatedly. (This distinguishes genuine dialectics from the liar's antinomy where in fact $A = \sim T$ holds.)

Many examples have been given. Here one more example from everyday life: driving an automobile.

Example 12.

T: steering an automobile
A: skidding (on ice, etc.) (negation)
S: countersteering to restore the course (negation of negation).

"Negation of negation" does not simply restore the original situation but implies *progress* (improved driving skill as a result of experience). We have essentially the phenomenon thesis – antithesis (negation) – synthesis (negation of negation). The progress is represented in Fig. 2.10 by the vertical distance $T - S$.

This law, negation of negation, may be considered the logical basis of progress or evolution in human life, human history, and human science. Contradiction between theory and experience (negation) leads to improved theories (negation of negation). Even Darwin's "survival of the fittest" also fits into our category. A danger (negation), successfully overcome (negation of negation), strengthens the personality as we have already remarked above.

(III) *Inversion of perspective*, "*dialectic reversal*". The best–known example is the Copernican revolution: placing the center of the universe at the sun (heliocentric system) instead of having the Earth in a central position (geocentric system) essentially simplified the law of planetary motion and led to Kepler's laws and Newton's mechanics. (The synthesis lies on almost too high a level: according to Einstein's general theory of relativity (sec. 3.4), all reference systems and all origins are theoretically (!) equivalent, cf. Example 3 on p. 46.)

By directing the attention of the observer from the object to his own perceptual apparatus (cf. sec. 1.4), Kant claimed to have provided a Copernican revolution in philosophy. In fact, it was followed by a period of intense flourishing of philosophy (Fichte, Schelling, Hegel).

By looking at the antinomy of the liar from an unexpected angle, Gödel was able to prove extremely profound and far–reaching theorems in logic (sec. 2.3). An *unprovable* proposition is recognized to be *true*.

A mathematical example which is less trivial than it looks, providing important generalizations, is as follows.

$\sqrt{2}$ is not a rational number; therefore it was not considered a number by the ancient Greeks. It can, however, be approximated with arbitrary accuracy by the rational numbers

$$\begin{aligned} x_1 &= 1.4 \\ x_2 &= 1.41 \\ x_3 &= 1.414 \\ x_4 &= 1.4142 \\ &\vdots \end{aligned} \tag{2.18}$$

The sequence x_1, x_2, x_3, ... does not have a rational number as a limit, it "leads to nowhere". The solution is Cantor's definition of $\sqrt{2}$ as precisely the *entire sequence* $\{x_1, x_2, x_3, \ldots\}$! This principle is frequently used in mathematics, cf. formula (2.8) on p. 27.

In adjustment theory (secs. 2.4, 2.6, and 4.4), the troublesome and annoying measuring errors are re-interpreted as stochastic variables with an interesting mathematical theory.

Mathematically "ill–posed problems" have become central in the recently fashionable "chaos theory" (sec. 3.2) and "inverse problems" (sec. 3.8)

Illness, pain, and grief if regarded from the proper point of view, may be recognized as positive factors in the development of human personality, "making a virtue of necessity".

In art, this is the principle of tragedy.

In philosophy, finally, the contrast between materialism (matter is primary, mind is derived from matter) and idealism (mind is primary, matter is derived) is, to a large extent, resolved by inverting the perspective, cf. sec. 5.3.

(IV) *Passage of quantity into quality*, and conversely. A classical example is the boiling or freezing of water, but any *phase transition* in physics or chemistry exemplifies this principle.

Such a phase transition is also the *emergence* of life from inorganic matter, and the transition from animals to man. The phase transitions are discontinuous (qualitative) but may be reached continuously (quantitatively). An example from geometry may illustrate that such a behavior is indeed possible (Fig. 2.13). As a matter of fact, this is just an illustration and by no means a proof!

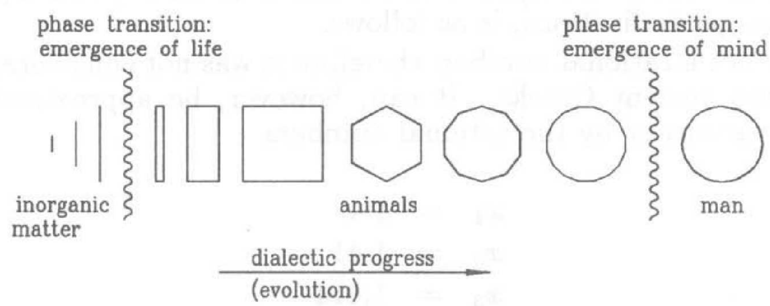


Figure 2.13: Illustration for the emergence of life (symbolized by various polygons) from inorganic matter (various lines), and of the emergence of the human mind (circle)

A rectangle, however narrow, is qualitatively different from a line, and a circle is different from a regular polygon, however small its sides and however large their number may be. Nevertheless, all transitions are continuous!

Does it surprise you that Cohen and Stewart (1994, p. 436–441) have linked emergence with Gödel's theorem?

This dialectical theory of phase transitions seems to be particularly attractive to natural scientists; cf. (Haldane 1939, p. 26). The excellent theoretical physicist Kaku (1994, p. 210) writes:

> This is the essence of dialectics. According to this philosophy, all objects ... go through a series of stages. Each stage is characterized by a conflict between two opposing forces ... When the conflict is resolved, the object goes to a higher stage, called the synthesis, where a new contradiction begins, and the process starts over again at a higher level. Philosophers call this the transition from "quantity" to "quality".

Dialectics and logic

Formal logic presupposes exactly defined concepts, such as tree, house, green: concepts that denote distinct objects or properties that do not change with time. It is impossible, however, to rigorously separate a tree or a house from its environment, and the color "green" is rarely found in pure form. Hence, exact concepts are idealizations, as we have already frequently seen. For such idealizations the laws of formal logic such as

$$\begin{array}{ll} \text{law of contradiction} & \sim (p \wedge \sim p) \\ \text{double negation} & \sim (\sim p) = p \end{array} \tag{2.19}$$

hold.

If dialectics is to apply to the real world, it must be wider than formal logic, somewhat like the "fuzzy logic" of sec. 2.4. Then these laws may not hold, as the examples of sec. 2.4 and Examples 7 to 11 of the present section have shown. Thus we may say:

Dialectic logic is a logic of rational approximation.

This is not the whole story, however. Take, for instance, the process of acquiring knowledge. According to Fichte, the "I" (the subject, T) confronts the "non–I", the surrounding world (the object, A). The "I" continuously takes some information from the surrounding world and thus gradually increases its knowledge through a continuous sequence of syntheses (S).

In this process, the "I" of one year ago (or one minute ago), is identical and yet not identical to the "I" now since, in the meantime, I have undergone a development. The law of contradiction no longer holds.

This important point was very well elaborated by Havemann (1964, pp. 48-49). Thinking begins with the dialectic (not formal–logic!) contradiction between *identity* and *difference*. Throughout the years we change: I am certainly different from what I was 50 years ago, nevertheless I feel that somehow I have retained my personal identity, and also legally I have definitely remained the same person.

Thus we may also say:

Dialectic logic is a logic of temporal evolution.

In several places above we have seen the importance of the concept of complementarity due to Niels Bohr. Thus we say:

Dialectic logic is the logic of complementarity.

Another aspect is the following. Formal logic only deals with well–defined and distinct *objects*. If we want to think about our own thinking, as we have remarked at the very beginning of the present section, in the context of a re-interpretation of Gödel's theorem, then the subject enters and logical paradoxes and antinomies may occur, as exemplified by the paradox of the liar (sec. 2.3). There arise logical structures which are beyond the reach of formal logic, e.g. "*The thinking thinks the thinking*" (Plotinus, A.D. 204–270). Subject, predicate, and object are identical in this sentence, yet it is not meaningless. Thus we may say:

Dialectic thinking is thinking about thinking.

Returning to sec. 2.3, we see that these are structures of *self–reference*, or *reflexive structures*. The well–known book (Hofstadter 1979) is full of such structures. Thus we may finally say:

Dialectic logic is reflexive logic.

In his 3 volume–work "The Science of Logic", Hegel proceeds from synthesis to synthesis, from level to level. His culminating chapter is entitled "The absolute idea". This crowning concept implicitly contains Hegel's whole system and, in principle, permits to derive it. This is somewhat similar to $\sqrt{2}$ "containing" all subsequent approximations 1.4, 1.41, 1.414, 1.4142, etc.

This example shows that Hegel's claim is not so extravagant as it looks. It really seems that *Hegel's logic is capable of deriving itself*.

To see what this means, consider an all–comprehensive scientific theory, fashionable under the term TOE, "theory of everything". If TOE is to explain "everything", *it also must explain itself*!

In a sense, Hegel's logic thus is the first TOE. As a matter of fact, much of Hegel's reasoning is "fuzzy" and difficult to follow. Nevertheless, if, some time in the future, a true TOE will be developed (which I doubt), it must, in a way, have incorporated Hegel's logic, hence it must admit self–reference. More about this will be said in sec. 6.6.

Bootstrapping. "Deriving itself" is strongly reminiscent of "bootstrapping", pulling oneself up by one's own bootstraps to get out of a swamp, like the famous Baron Munchausen did. Learning a foreign language on the basis of one's own language may need considerable effort but is simple "in principle". A child learning his/her first language, however, is performing a remarkable feat of bootstrapping, learning "more language" on the basis of "some language" already acquired (Hofstadter 1979, p. 294). Bootstrapping is now also popular in elementary particle physics, cf. sec. 6.6.

Dialectics and symbolic logic

Dialectic logic operates on several "levels", cf. Fig. 2.8 on p. 46, whereas symbolic logic remains always on the same level. A classical example of this is Gödel's proof mentioned at the beginning of this section.

Attempts to formalize dialectic logic have failed. Thus dialectic thinking cannot be implemented on a computer. As we have seen at the beginning, dialectic logic rather is a higher–level informal commentary. Still, a geometric interpretation is possible; see "The geometry of complementarity" in sec. 2.6 below.

A deplorable source of misunderstanding has been the *dialectic contradiction*. It is, of course, not the straight negation of symbolic logic, but something much more subtle (for a simple model see "The geometry of complementarity" below). Dialectic thinking seems to be a habit of thinking that must be practiced, like mathematics. Some of the greatest formal logicians (Gödel, Whitehead, Smullyan) at the same time are superb dialectic thinkers. We also mention the mathematician Andreas Speiser (1952, 1955, 1959) and the physicist Niels Bohr (1934, 1958, 1963); cf. also (Folse 1985).

2.6 Geometry: dimensions two to infinity

Ὁ θεὸς ἀεὶ γεωμετρεῖ.

Plato

Euclidean geometry has been a prototype of a rigorous axiomatic system, from its discovery by Euclid around 300 B.C. to the present day. The great philosopher Spinoza in the 17th century formulated his philosophical system in the style of Euclid, with definitions, axioms, and theorems. Until a hundred years or so ago, geometry was taught in schools according to Euclid's elements. Small wonder if boys and girls got to hate Euclid. — Only recently, David Hilbert improved upon Euclid by giving a fully rigorous axiomatic system of Euclidean geometry in the modern sense of sec. 2.2.

The father of Greek geometry was the half–mythical figure of Pythagoras (around 530 B.C.). He founded a philosophical school with a mystical theory about numbers: "Number is the essence of things". This sounds surprisingly modern: in fact, contemporary physics is thoroughly mathematical. He did not find the *theorem of Pythagoras*,

$$c^2 = a^2 + b^2 \quad , \tag{2.20}$$

between the three sides a, b, c of a rectangular triangle (Fig. 2.14). This *fact* was known much earlier, but Pythagoras gave the first exact *proof*.

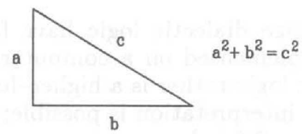


Figure 2.14: The theorem of Pythagoras

In ancient Greece, "geometry" was more or less synonymous with "mathematics". The greatest philosopher, not only of Greece but of all times, Plato (428–347 B.C.) had a particularly high regard of mathematics. His school was called "Academy", and the entrance of the Academy showed the inscription "No one ignorant of geometry may enter here". When one of his students asked what the occupation of God was, he answered "God always geometrizes" (in Greek the motto above the present section). Only dialectics is higher: the mathematicians go out hunting, and the dialecticians use their catch (Speiser 1955, p. 61).

Other Greek mathematicians such as Archimedes made great discoveries too. In the Middle Ages there was hardly any progress in mathematics or geometry.

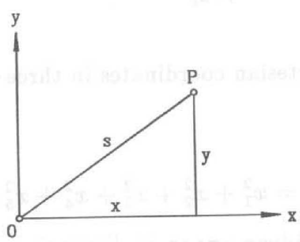


Figure 2.15: Cartesian coordinates in the plane

A basic new discovery was made by the philosopher René Descartes (Cartesius, 1596–1650): the *Cartesian coordinates*, which permitted to treat geometrical problems by algebraic methods. A point P in the plane may be defined by its coordinates x and y referred to a rectangular coordinate system xOy (Fig. 2.15). The distance s of P from the origin O is given, according to Pythagoras' theorem:

$$s^2 = x^2 + y^2 \quad . \tag{2.21}$$

In our usual three–dimensional space (*Euclidean space*) we have three coordinates, called x, y, z or x_1, x_2, x_3 (Fig. 2.16) and the distance s becomes

$$s^2 = x^2 + y^2 + z^2 = x_1^2 + x_2^2 + x_3^2 \quad . \tag{2.22}$$

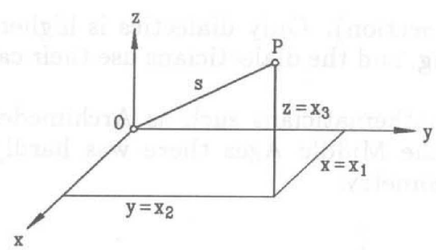


Figure 2.16: Cartesian coordinates in three–dimensional space

We can visualize two–dimensional space (the plane) and three–dimensional space. A space of dimension 4 is space–time with coordinates x, y, z, t; t denotes time. Space of higher dimension cannot be visualized, but the Cartesian method makes them accessible to an algebraic treatment which is not essentially more difficult than for dimensions 2 or 3.

For dimension 4 we have

$$s^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 \quad , \tag{2.23}$$

for dimension 5

$$s^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 \quad , \tag{2.24}$$

and generally for Euclidean space of dimension n,

$$s^2 = x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2 \quad , \tag{2.25}$$

generalizing (2.21) and (2.22).

Such n–dimensional spaces are of fundamental importance as *auxiliary concepts* in physics and in other natural sciences.

We may let n become greater and greater: $n = 10$, 100, 1000, a million, a billion, etc. We may even let n become infinitely great, we let n go to infinity, in mathematical symbols $n \to \infty$. Then (2.25) may be generalized to

$$s^2 = x_1^2 + x_2^2 + x_3^2 + \cdots \text{ to infinity} \quad . \tag{2.26}$$

A necessary and sufficient condition is that the infinite sum (2.26) "converges", as mathematicians say, that is, that s results as a finite number. This is *Hilbert space*; it is the generalization of n–dimensional Euclidean space for $n \to \infty$.

Hilbert space is *the* mathematical tool for quantum mechanics, as we shall see in sec. 3.5.

Spaces of higher dimension and Hilbert space appear awesome concepts for the non–initiated. The main (or even only) problem is psychological. The psychological barrier is best overcome by practical computations. For instance, solving a linear system of equations with 5 unknowns means that we work in *5–dimensional Euclidean space*. And if your pocket calculator solves a system with 10 unknowns, you work in 10–dimensional space! What could be easier than that?

"But I cannot visualize these spaces!", you say. Let me tell you a secret. Nobody can visualize a higher–dimensional space, but all mathematicians speak fluently about points, straight lines, planes, or subspaces in n–dimensional space. The picture they have in mind when they speak in this way, is *our usual three–dimensional space* (or even the two–dimensional plane). Computing in n–space and at the same time visualizing 3–space may sound a little schizophrenic but it is always done in this way, and it always works! So don't worry. This schizophrenic way of visualization even works in infinite–dimensional Hilbert space, as we shall see in sec. 3.5.

Differential calculus. It was invented simultaneously and independently by the great philosopher Leibniz (1646–1716) and by the perhaps even greater physicist Newton (1642–1727). This led to a terrible quarrel about priority: even great personalities are only human beings.

What we need here is extremely simple: the Pythagorean theorem (2.21) for infinitely small (*infinitesimal*) differences dx, dy, and ds (Fig. 2.17); ds is called line element. We obviously have

$$ds^2 = dx^2 + dy^2 \quad . \tag{2.27}$$

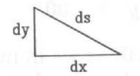


Figure 2.17: The Pythagorean theorem for an infinitesimal triangle

What means "infinitesimal"? Somewhat loosely speaking, it means "very small", "as small as you like", the smaller the better. For some purposes, 1 meter may be "infinitesimal" (for instance in a map); if you don't like this, take 1 centimeter, 1 millimeter, or anything as small you like. (Mathematicians will give you a more precise definition.)

Similarly, we have in three dimensions,

$$ds^2 = dx^2 + dy^2 + dz^2 \tag{2.28}$$

and in n dimensions

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + \cdots + dx_n^2 \quad . \tag{2.29}$$

At this point, you may say, somewhat disappointedly, "But this is quite easy!". Dear reader, it *is* easy, believe me and don't waste your time looking for complications.

The following application is slightly more demanding mathematically. You can, however, safely disregard the formulas if you wish.

Geometry of least–squares adjustment

Consider a system of linear equations

$$\begin{aligned} a_{11}x + a_{12}y &= l_1 \quad , \\ a_{21}x + a_{22}y &= l_2 \quad . \end{aligned} \tag{2.30}$$

The coefficients a_{ij} are known, the left–hand sides l_1 and l_2 have been measured, and x and y are unknowns to be determined. For clarification, let me give a numerical example:

$$\begin{aligned} 2x - 3y &= 10 \quad , \\ x + y &= 30 \quad . \end{aligned} \tag{2.31}$$

The solution is $x = 20$, $y = 10$, as we immediately find on substitution.

It frequently happens that more observations l_i are made, in order to check the determination and to improve its numerical value. For instance, let a third observation $l_3 = 8$ be made, which gives the system

$$\begin{aligned} 2x - 3y &= 10 \quad , \\ x + y &= 30 \quad , \\ x - y &= 8 \quad . \end{aligned} \tag{2.32}$$

The solution of the second and third equations gives $x = 19$, $y = 11$, which obviously is different from the solution of the first and second equations, namely $x = 20$ and $y = 10$. We say that the three equations (2.32) are *inconsistent*.

Let

$$\begin{aligned} a_{11}x + a_{12}y &= l_1 \quad , \\ a_{21}x + a_{22}y &= l_2 \quad , \\ a_{31}x + a_{32}y &= l_3 \quad . \end{aligned} \tag{2.33}$$

be such an inconsistent system. To find a solution, we must slightly change the l_i, obtaining

$$\begin{aligned} l_1 + v_1 &= a_{11}x + a_{12}y \quad , \\ l_2 + v_2 &= a_{21}x + a_{22}y \quad , \\ l_3 + v_3 &= a_{31}x + a_{32}y \quad . \end{aligned} \tag{2.34}$$

We now introduce the very reasonable principle that the changes v_1, v_2, v_3 should be as small as possible, more precisely

$$v_1^2 + v_2^2 + v_3^2 \Longrightarrow \text{minimum} \quad . \tag{2.35}$$

This is best written in vector–matrix symbolism which should be understood immediately. Equation (2.34) thus becomes

$$\mathbf{l} + \mathbf{v} = \mathbf{A}\,\mathbf{x} \quad , \tag{2.36}$$

where

$$\mathbf{l} = \begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix} \quad , \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \quad , \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} \tag{2.37}$$

are *vectors* and

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \tag{2.38}$$

is a rectangular *matrix*. The condition (2.35) is thus abbreviated as

$$\mathbf{v}^T\mathbf{v} \Longrightarrow \text{minimum} \quad , \tag{2.39}$$

where

$$\mathbf{v}^T = [v_1 \ v_2 \ v_3] \tag{2.40}$$

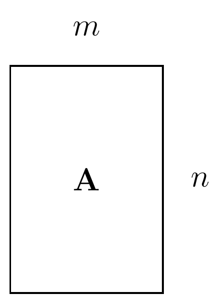
is the *transpose* of the vector $\mathbf{v}$ in (2.37).

The solution can be written in the matrix form

$$\mathbf{x} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{l} \quad , \tag{2.41}$$

where $\mathbf{M}^{-1}$ denotes the inverse of a "square" matrix $\mathbf{M}$.

The advantage of this symbolic notation is that $\mathbf{l}$ can be any n–vector, $\mathbf{x}$ some m–vector, and $\mathbf{A}$ a matrix of n "rows" and m "columns". The only condition imposed on the integers m and n is that $m < n$, so that we have an inconsistent system with a "standing" matrix $\mathbf{A}$:



What is interesting here is that the solution (2.41) of the problem defined by (2.36) and (2.39) is a projection of a vector $\mathbf{l}$ in n–dimensional space onto an m–dimensional subspace.

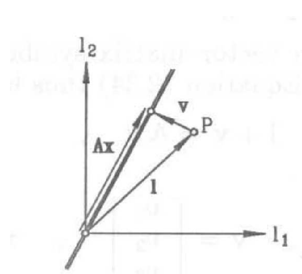


Figure 2.18: Adjustment as projection

Terrible, isn't it? Fig. 2.18 immediately clarifies the situation. If the system of equation were consistent, then $\mathbf{l}$ would lie in the subspace symbolized in Fig. 2.18 by a double line. Since they are inconsistent, $\mathbf{l}$ lies a little off the subspace and must be projected onto it by adding the vector $\mathbf{v}$. Condition (2.35) requires that the vector $\mathbf{v}$ represents the *shortest distance* of the point P from the subspace (line, plane, etc.). Thus $\mathbf{v}$ must be *orthogonal* to the subspace.

The determination of the unknown vector $\mathbf{x}$ by the condition $\mathbf{v}^T\mathbf{v} \Longrightarrow$ minimum is called *least–squares adjustment*. "Least–squares" because the square sum of the "corrections" v_i ($i = 1, 2, \ldots, n$) must be a minimum, cf. (2.35). "Adjustment" means that we must adjust the data l_i slightly and "democratically" (equal treatment for v_1, v_2, v_3!) in order to fit the system. See the example of eqs. (2.13) and (2.14) on p. 41.

Note that we have illustrated a projection of a vector in n–dimensional space onto a m–dimensional subspace by the simple diagram of Fig. 2.18 where $n = 2$ and $m = 1$. This is simple and represents all the essential geometry. There is no need to visualize a general n–dimensional space which nobody can do anyway.

It is like in a discussion in which all participants have slightly different opinions. In order to reach general agreement, all opinions (l_i) must be slightly "adjusted", everyone has to make some concessions (v_i!). It is, however, desirable that these concessions should be as slight as possible ($\mathbf{v}^T\mathbf{v} \Longrightarrow$ minimum).

Remember this principle of adjustment! We shall meet it again. If you could not quite follow our mathematical argument, never mind. Just remember the comparison with a discussion and the "democratic adjustment" of the participants' opinions.

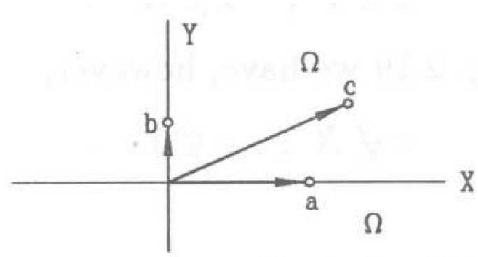


Figure 2.19: The plane Ω as the "union" of two orthogonal straight lines (subspaces) X and Y.

The geometry of complementarity

(This is only for the courageous. If you find it too difficult, simply skip it.) Geometry gives an ideal model to understand dialectic opposition (complementarity) and ordinary logical contradiction. Orthogonal subspaces are also called *complementary*. Consider two mutually orthogonal straight lines X and Y (Fig. 2.19). They span the plane in which our figure lies. We say that both straight lines X and Y are complementary (mutally orthogonal) subspaces of the plane Ω, or that Ω is the generalized *union* of X and Y:

$$\Omega = X \cup Y \quad .$$

Because of the orthogonality, the (generalized) *intersection* of X and Y is zero:

$$X \cap Y = \emptyset \quad .$$

Now we write

$$\begin{aligned} &X \subset \Omega\ , \quad Y \subset \Omega \\ &Y = \bar{X} \text{ (not } X\text{)}, X = \bar{Y} \text{ (not } Y\text{)} \end{aligned}$$

to indicate that the lines X and Y are complementary subspaces of Ω. Now, clearly, the vector a lies in the plane:

$$a \in \Omega \quad ,$$

and it lies in the subspace X:

$$a \in X\ , \quad a \notin Y$$

and similarly for the vector b:

$$b \in Y\ , \quad b \notin X \quad .$$

For the vector c in Fig. 2.19 we have, however,

$$c \notin X\ , \quad c \notin Y \quad .$$

Let us compare this situation with the set theory of ordinary logic (sec. 2.1). Here Ω is a set (represented by a square) which is decomposed

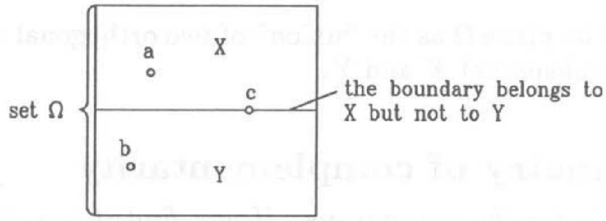


Figure 2.20: The set Ω as the union of the two sets X and Y.

into two subsets (rectangles X and Y (Fig. 2.20). Again we have

$$\Omega = X \cup Y$$
$$X \cap Y = \emptyset$$
$$X \subset \Omega$$
$$Y \subset \Omega$$
$$a \in X\,, \quad a \notin Y$$
$$b \in Y\,, \quad b \notin X$$

as before, but *now*

$$c \notin X\ , \quad c \notin Y = \bar{X}$$

is impossible. The straight boundary is considered to belong to the set X but not to Y. Hence even for a point c on the boundary,

$$c \in X\ , \quad c \notin Y \quad .$$

Denoting

$$\begin{aligned} \Omega - X &= Y = \bar{X} \text{ (not } X\text{)} \\ \Omega - Y &= X = \bar{Y} \text{ (not } Y\text{)} \end{aligned}$$

we have the mutually exclusive alternatives

$$\begin{gathered} c \in X \quad \text{or} \quad c \in \bar{X} \\ p \quad \text{or} \quad \sim p \\ p \vee \sim p \quad , \end{gathered}$$

which is the law of the excluded middle, eq. (2.11) on p. 29.

The situation of Fig. 2.19 is essentially different, although we have purposely used formally identical notations. Here we have

$$c \notin X \ , \quad c \notin Y = \bar{X} \quad ,$$

so that the law of the excluded middle

$$\text{either } c \in X \quad \text{or} \quad c \in \bar{X} \quad (p \vee \sim p)$$

does not hold.

Thus the logic of *complementary subspaces* (Fig. 2.19) is essentially different from the ordinary logic of *complementary subsets* of sec. 2.1 and Fig. 2.20.

In fact, Fig. 2.19 is for *quantum logic* what Fig. 2.20 is for ordinary logic.

Fig. 2.19 contains, in a nutshell, all the strange quantum phenomena of complementarity; it is the simplest example for a dialectic logic. Thus dialectic complementarity, $Y = \bar{X}$ according to Fig. 2.19, is essentially different from "ordinary" logical negation, $Y = \bar{X}$ according to Fig. 2.20. Thus the "dialectic complement" of the X–axis is the Y–axis, whereas the "ordinary" complement of the X–axis would be the whole plane excluding the X–axis. Actually, in quantum theory (sec. 3.5), the plane of Fig. 2.19 should be an infinite–dimensional Hilbert space, but we do not want to misuse the patience of the reader who has kindly followed us so far.

The main purpose of this example was to show that dialectic negation is meaningful. Here Hegel (dialectics), Bohr (complementarity), and Weizsäcker (quantum logic) meet.

Part B

Natural Science

Chapter 3

Physics

3.1 Classical mechanics and determinism

Nature and Nature's Laws lay hid in Night:
God said, Let Newton be! and All was Light.

Alexander Pope

The great Greek philosopher Aristotle (384–322 B.C.) believed that the *velocity* of a body is proportional to the force to which it was subjected. Ordinary experience seems to confirm this view. A horse carriage moves the faster, the stronger the horses are. A body lying on the floor does not move unless some force is exerted to drag it along.

Only Galileo Galilei (1564–1642) recognized that matters are not so simple. A body lying on a very smooth and plane ice surface will continue to move with constant velocity and in a constant direction even if the initial force has ceased to act. To be sure, this body will gradually slow down and finally stop, but the cause is *friction*. If there is no friction, the movement will be continuous and will never come to a stop. A space ship in intergalactic space will forever move with constant speed along a straight line after the rocket engines have been shut off. Thus Aristotle and common sense have been deceived by friction.

The correct law of motion in the absence of friction was discovered by Isaac Newton (1642–1727). It has the form

$$m\ddot{\mathbf{x}} = \mathbf{F} \quad . \tag{3.1}$$

Here m denotes the mass, and $\mathbf{F}$ is the force. The position vector is

$$\mathbf{x} = [x,\, y,\, z] \quad , \tag{3.2}$$

the velocity vector is its time derivative:

$$\dot{\mathbf{x}} \equiv \frac{d\mathbf{x}}{dt} = [\dot{x},\, \dot{y},\, \dot{z}] \quad , \tag{3.3}$$

and the acceleration is the second derivative:

$$\ddot{\mathbf{x}} \equiv \frac{d^2\mathbf{x}}{dt^2} = [\ddot{x},\, \ddot{y},\, \ddot{z}] \quad . \tag{3.4}$$

Thus *Newton's law of motion* (3.1) says that the *acceleration* is proportional to the force, and not the velocity as Aristotle thought.

In order to fully define the movement, in addition to the *differential equation* (3.1) we need *initial conditions*: at a certain instant $t = t_0$, the position and velocity:

$$\mathbf{x}_0 = \mathbf{x}(t_0) \quad , \qquad \dot{\mathbf{x}}_0 = \dot{\mathbf{x}}(t_0) \tag{3.5}$$

must be given.

Assume motion under no force, $\mathbf{F} = 0$. Then (3.1) gives

$$\ddot{\mathbf{x}} = 0 \quad . \tag{3.6}$$

The solution of this differential equation is

$$\mathbf{x} = \mathbf{a}\, t + \mathbf{b} \quad , \tag{3.7}$$

where the constant vectors **a** and **b** serve as integration constants. To understand this, differentiate (3.7) twice:

$$\dot{\mathbf{x}} = \mathbf{a} \quad , \tag{3.8}$$

$$\ddot{\mathbf{x}} = 0 \quad . \tag{3.9}$$

Thus (3.6) is satisfied, what was to be shown. If we put $t = 0$ in (3.7) and (3.8), we get

$$\mathbf{a} = \dot{\mathbf{x}}_0 \quad , \qquad \mathbf{b} = \mathbf{x}_0 \quad . \tag{3.10}$$

Taking for the initial instant $t_0 = 0$, we thus have a very instructive interpretation of the integration constants **a** and **b**: they are nothing else than the initial conditions (3.5).

Newton's law of gravitation. Besides Newton's law of motion (3.1), we also have his law of gravitation:

$$F = G\frac{m_1\, m_2}{l^2} \quad . \tag{3.11}$$

Two point masses (Fig. 3.1) attract each other with a force **F** of magnitude F, proportional to the masses m_1 and m_2, and inversely proportional to the square of their distance l; this is the famous *inverse square law*. Here G denotes a universal constant, the *gravitational constant*. We also have the equality of *action and reaction*: the two forces $\mathbf{F}_1$ and $\mathbf{F}_2$ in Fig. 3.1 are equal in magnitude and opposite in direction. The magnitude of both $\mathbf{F}_1$ and $\mathbf{F}_2$ is given by (3.11).

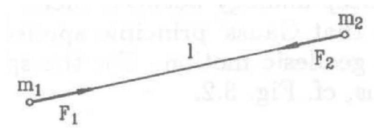


Figure 3.1: Illustrating the law of gravitation

If Newton's law of gravitation (3.11) is used in the equation of motion (3.1), then this differential equation, on integration, gives the Kepler ellipses, along which the planets move around the Sun.

Principles of mechanics. If the motion is subject to constraints, the simple Newtonian equation of motion is no longer applicable. For instance, frictionless motion of a particle constrained to move along a curved surface cannot be along a straight line, even if there is no external force, $\mathbf{F} = 0$. The "straightest" curve on a surface is a *geodesic*, representing the shortest line between two points that wholly lies in the surface. If the surface is a sphere, then the geodesic is a great circle. Now it can be shown that *frictionless and forceless motion along a surface really is motion with constant velocity along a geodesic.* Even this simple but important case is not covered by (3.1).

So for motion on a surface Newton's equation (3.1) is not satisfied, that is,

$$m\ddot{\mathbf{x}} - \mathbf{F} \neq 0 \quad . \tag{3.12}$$

If the left–hand side cannot be zero, then let us try at least to make it as small as possible:

$$(m\ddot{\mathbf{x}} - \mathbf{F})^2 \Longrightarrow \text{minimum} \quad , \tag{3.13}$$

subject to the given conditions, for instance, motion on a surface. This is *Gauss' principle of least constraint.*

It is in full analogy to the *principle of least squares* discussed in sec. 2.6, eq. (2.35) on p. 64. In fact, (2.36) says that

$$\mathbf{A}\,\mathbf{x} - \mathbf{l} \neq 0 \quad , \tag{3.14}$$

and (2.39) is equivalent to

$$(\mathbf{A}\,\mathbf{x} - \mathbf{l})^2 \Longrightarrow \text{minimum} \quad . \tag{3.15}$$

The analogy between (3.12) and (3.13), on the one hand, and (3.14) and (3.15), on the other hand, is obvious.

Thus it is not surprising that both principles are due to Gauss, who also recognized the deep analogy between them.

It may be shown that Gauss' principle applied to a free particle on a surface, does give geodesic motion. For the sphere, motion along a great circle is obvious, cf. Fig. 3.2.

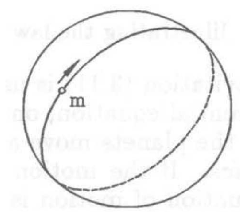


Figure 3.2: A free particle describes a geodesic on a sphere

For many other simple and complicated cases, Newton's elementary law (3.1) does not directly apply. A pertinent example is the rotation of a rigid body, because Newton's equations are essentially valid for point masses only and do not apply to rotation. With respect to orbital motion about the Sun, the planets may be considered point masses, but Earth rotation must be treated in a different way.

A number of other principles, more general than Newton's laws, were proposed in the 18th century by d'Alembert, Lagrange and others. This is subject of *analytical dynamics*, of which a non–specialist account can be found in (Lindsay and Margenau 1957, Chapter III). We have briefly considered only Gauss' principle and shall now outline Hamilton's method.

Hamilton's equations. The Newton equation (3.1) is in reality a system of *three* ordinary differential equations *of second order*:

$$\begin{aligned} m\ddot{x}_1 &= F_1(x_1, x_2, x_3) \quad , \\ m\ddot{x}_2 &= F_2(x_1, x_2, x_3) \quad , \\ m\ddot{x}_3 &= F_3(x_1, x_2, x_3) \quad , \end{aligned} \tag{3.16}$$

$x_1 = x$, $x_2 = y$, $x_3 = z$ denoting the Cartesian coordinates which are the components of the position vector $\mathbf{x}$, and similarly F_1, F_2, F_3 for the force vector $\mathbf{F}$.

Now we introduce the auxiliary quantities

$$p_1 = m\dot{x}_1 \ , \quad p_2 = m\dot{x}_2 \ , \quad p_3 = m\dot{x}_3 \tag{3.17}$$

called *momenta*. The coordinates x_1, x_2, x_3 are now denoted by q_1, q_2, q_3. Then (3.17) and (3.16) become with $i = 1, 2, 3$:

$$\begin{aligned} \dot{q}_i &= \frac{1}{m} p_i \quad , \\ \dot{p}_i &= F_i(q_1, q_2, q_3) \quad . \end{aligned} \tag{3.18}$$

Thus we have reduced the *three* differential equations (3.16) of *second* order by $3 + 3 =$ *six* differential equations of *first* order.

This method is standard in the theory of differential equations and not particularly enlightening.

What is significant, however, is the fact that William Hamilton (1788–1856) was able to bring (3.18) to the form

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \quad , \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i} \quad , \end{aligned} \tag{3.19}$$

with *one* function H only, instead of the 3 functions F_i! Because of their importance, they are called the *canonical equations* of mechanics, and H is known as Hamilton's function or, briefly, as *Hamiltonian*. By the way, H is simply the sum of kinetic and potential energy. Any quantities p_i and q_i satisfying (3.19) are called *canonically conjugate variables*.

The true importance of the Hamiltonian equations (3.19), however, is the fact that q_i need not be Cartesian coordinates but can be any *generalized coordinates* (parameters), and i need not be restricted to 1, 2, 3 but can assume so many values as we need parameters to fully describe the dynamical system. For instance, for a rotating rigid body we need 6 parameters q_i ($i = 1, 2, 3, 4, 5, 6$): three translations (along the x, y, z axes) and three rotations (e.g., around the same axes). If we have r particles, then we need $3r$ parameters q_i: 3 for each particle.

Let us assume that we have n generalized coordinates q_i. Then we have $2n$ differential equations (3.19), and we can solve them uniquely provided we have the $2n$ *initial values* q_i and p_i at time $t = t_0$.

Laplace's demon.

An intelligent being which, for some given moment of time, knew all the forces by which nature is driven, and the relative position of the objects by which it is composed (provided the being's intelligence were so vast as to be able to analyze all the data), would be able to comprise, in a single formula, the movements of the largest bodies in the universe and those of the lightest atom: nothing would be uncertain to it, and both the future and the past would be present to its eyes. The human mind offers in the perfection which it has been able to give to astronomy, a feeble inkling of such an intelligence.

This impressive statement was given by Pierre Simon de Laplace (1749–1827); the "intelligent being" has become famous as "Laplace's demon".

This is the classical expression of *causality* or *determinism*: given the equations of motion and the initial conditions at $t = t_0$, the state of the system is exactly known at all earlier ($t < t_0$) and all later ($t > t_0$) times. Determinism reigned supreme until about 1925, when quantum theory started thoroughly to shake it (sec. 3.5).

Recently, however, determinism has come under attack even from its very stronghold, classical mechanics. This has been achieved by the theory of chaotic systems (sec. 3.2).

The principle of least action. Instead of differential equations, classical mechanics can also be expressed by an *integral* minimum principle of form

$$\int_A^B L\,dt \Longrightarrow \text{minimum} \tag{3.20}$$

where an integral (the "action") of a function L is to be minimized. The *Lagrangian* L is related to the energy and also to the Hamiltonian H in a way which is not necessary for the present argument. Least–action principles have been given by several scientists starting with Pierre Louis de Maupertuis (1698–1759) and Leonhard Euler (1707–1783).

From the *integral* principle (3.20) it is possible uniquely to derive the *differential* equations (3.19). This is of considerable philosophical importance, for the following reasons.

An integral principle (3.20), minimizing (or maximizing) some "overall" quantity, has been interpreted as expressing a tendency of nature towards perfection, attaining some ideal: maximum or optimum sounds better than minimum, but is essentially the same thing. It thus expresses a *finalist tendency*, a "causa finalis" in the sense of Aristotle, cf. sec. 5.4. Such finalism occurs especially in biology (sec. 4.1). It has been opposed to the causal determinism as exemplified by the differential equations of classical mechanics.

The deduction of the *deterministic* equations (3.19) from the *finalistic* integral (3.20) shows that both principles can coexist peacefully: the principle (3.20), so to speak, creates its own differential equations (3.19).

In a similar way we shall see in sec. 4.1 that a thermostat, governed by a "finalistic" principle of producing a desired temperature, will "generate" its own physical "deterministic" differential equations that help achieve the goal.

Thus causality, characteristic for classical mechanics, and finalism, considered typical for biology, are far less incompatible as they first appear, cf. also (Thom 1975, sec. 12.1.A).

"Causality", so to speak, is the answer to the question "For which reason?", whereas "finality" answers the question "For which purpose?".

The basic results of the present section will also be needed to discuss geodesic motion in general relativity (sec. 3.4) and a generalization of Hamiltonian methods to quantum theory (sec. 3.5).

But also taken in itself, classical mechanics has an incredibly rich structure. It comprises:

– causality: basic property;

– chaos: sec. 3.2;

– final causation: just discussed;

– constraints: eq. (3.13); and even

– "software laws" in a rudimentary form: as initial conditions (see also sec. 4.5).

Ideas will be needed rather than formulas, so the reader need not understand all mathematical details. Interested readers may consult any textbook on theoretical physics; particularly suited for the present purpose is the treatment in (Lindsay and Margenau 1957, Chapter III). We also mention (Margenau 1950) which is less mathematical and more philosophical and which is still a classic.

3.2 Deterministic chaos

In the beginning ... there was Poincaré.
E. Atlee Jackson

The deterministic paradise of classical mechanics, over which Laplace's demon (sec. 3.1) exerted a rigid but essentially benevolent, orderly, and stable regime, began to show, on closer inspection, some strange and irritating features.

The application of mechanics to gases and fluids consisting of an enormous number of particles (molecules) led to the statistical theory of heat. Heat was explained as the random and irregular, more or less violent motion of these particles. In view of the enormous number of these particles, it is practically impossible to describe the trajectory of every particle by Newton's laws (even assumed that this would be theoretically possible). Instead, these particles were treated statistically, which led to *statistical mechanics* or *statistical thermodynamics*, created by Josia Willard Gibbs (1839–1903), Ludwig Boltzmann (1844–1906) and others. A brilliant success was the derivation of the basic equations of thermodynamics from the principles of classical mechanics combined with statistical considerations. *Temperature* was explained in terms of the average kinetic energy of the molecules; it is the higher, the greater the average velocity of the particles is. The important concept of *entropy* was introduced, and Boltzmann found his famous equation, formula (4.3) of sec. 4.3.

But here a problem arises. The equations of classical mechanics are *time–reversible*. This means that these equations retain their form on replacing time t by $-t$. On the other hand, the equations of thermodynamics are typically *irreversible*: the entropy in a physical system always increases, see eq. (4.4). This contradiction must be due to the introduction of statistics, either because of the enormous amount of particles, or because of the incredibly complicated, "chaotic", shape of the trajectories of the particles (or both). These controverses, in which already Boltzmann was involved, led to very important advances in physics, mathematics, and probability theory (sec. 3.3), known by the name of *ergodic theory*.

The French mathematician Henri Poincaré (1854–1912) found already in 1890 that even relatively "simple" *nonlinear* dynamical problems in astronomy etc. may admit extremely complicated, irregular, even "chaotic" trajectories. In his classical work "Les Méthodes nouvelles de la Mécanique céleste" (1899) vol. III, p. 389 he wrote:

Imagine the figure formed by these two curves and their infinitely many intersections ...; these intersections form a kind of meshwork, tissue, or infinitely dense network ... One is struck by the complexity of this figure which I do not even attempt to draw. Nothing is better suited to give us an idea of the complexity of the three–body problem and in general of all the problems of dynamics in which there is no uniform integral [of the motion] ...

The modern theory of *general nonlinear dynamical systems* is considered to start with Poincaré's work. The subject then lay relatively dormant, known only to a few specialists, until 1954 when the famous Russian mathematician Andrei Kolmogorov (1903–1987) and his younger colleague Vladimir Arnold started with a general and systematic treatment of such strange trajectories. In 1963 there followed an independent paper on an application to meteorology by the American Edward Lorenz. Then the subject exploded. Currently it is probably the most popular subject of mathematics, known to a broad general public.

Let me try to explain what Lorenz did. He took the equations of mathematical *weather prediction*, simplified them and studied the solution numerically with the help of a computer. These solutions proved to be extremely *unstable*: two solutions with almost identical initial conditions started to diverge wildly (Fig. 3.3). Since the data of meteorology are unavoidably insufficient and inaccurate, the initial conditions are not exactly known; small deviations result in completely different behavior. This is the reason why it is hardly meaningful to make detailed weather predictions more than a few days ahead. (In astronomy, predictions are good for tens or even hundreds of years, in spite of Poincaré ...)

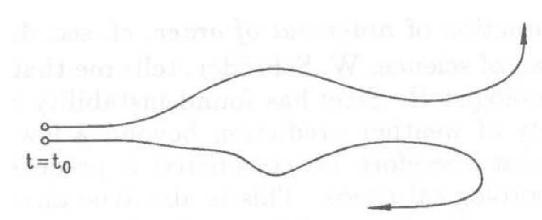


Figure 3.3: Two unstable trajectories

Let us repeat:

stability: small causes produce small effects;
instability: small causes produce large effects.

Classical causality implicitly presupposes stability. Stability is the environment in which Laplace's demon thrives.

Unstable systems are mathematically described always by *nonlinear* differential equations. Therefore, as we have already mentioned, mathematicians speak of general nonlinear dynamical systems. (Popularly speaking, the difference between "linear" and "nonlinear" is essentially the difference between a straight and a curved line; the function $y = 2 + 3x$ is linear, whereas the functions $y = x^2$ and $y = \sin x$ are nonlinear.) Unstable nonlinear dynamical systems are nowadays widely known by the name of *chaos theory*.

We distinguish between *conservative* dynamic systems for which the total energy is conserved (e.g., those described by Hamiltonian equations (3.19)), and *dissipative* systems for which part of the energy is dissipated as heat, e.g., through friction.

The nonlinear systems of *celestial mechanics* as investi gated by Poincaré, Kolmogorov, and Arnold are *conservative*. The *meteorologic systems* studied by Lorenz are *dissipative*, because the atmosphere constantly receives energy from the sun and radiates it again into outer space: otherwise "global warming" would be very rapid indeed. The name, *chaotic systems*, is particularly appropriate for meteorological and similar dynamic systems.

Chaos theory is an outstanding example of a theory as an instrument for discovery, a "searchlight": now chaotic phenomena are found everywhere, from clouds to earthquakes, and from turbulent mountain streams to human heartbeats. Deterministic chaos, so to speak, is an example of *chaos out of order*. There is also an emergence of *order out of chaos*; cf. the derivation of thermodynamics from statistical mechanics and sec. 3.3. In fact, both cases are closely interrelated and related also to the production of *order out of order*, cf. sec. 4.3 (p. 182).

The historian of science, W. Schröder, tells me that the well–known German meteorologist H. Ertel has found instability as the reason for the impossibility of weather prediction beyond a few days already in 1941. Ertel must therefore be considered a predecessor of Edward Lorenz in meteorological chaos. This is also true already for Poincarè (1908), as the quotation in sec. 6.3 (p. 243) shows.

Suggested additional reading. There is an incredible amount of books and papers on chaos theory. An advantage of its popularity is the fact that there are outstanding presentations for the general public, of high level but without formulas. An extremely readable introduction is (Gleick 1988); Stewart (1990) is a fascinating presentation of all the details but without formulas; an authoritative and very readable introduction is (Lorenz 1993); and Abraham and Shaw (1984) managed to present the intricate geometry, which was even too much for Poincaré as his quotation shows, in beautiful pictures which should be accessible to everyone with an interest in science. Applications to biology and medicine may be found in (Glass and Mackey 1988). Chaos theory is very popular also because its geometrical structures (fractals, strange attractors) are of a truly exotic beauty. Particularly remarkable is the combination of beauty and readability in (Briggs 1992). For statistical mechanics and thermodynamics and their philosophical implications, (Lindsay and Margenau 1957) is still unsurpassed.

3.3 Probability

God does not throw dice.

Albert Einstein

Nor is it our business to prescribe
to God how He should run the world.

Niels Bohr

A simple and extremely instructive example of an unstable motion is *throwing a die*. The die is supposed to be a perfect, absolutely homogeneous cube, whose faces are numbered 1, 2, 3, 4, 5, 6.

If we throw it, it will come to rest showing, say, face 3. If we throw it again, trying to repeat the first throw as accurately as possible, it may show a 6 (Fig. 3.4). The *initial conditions* defined by the way of throwing may be almost identical; nevertheless the results will be quite different and practically completely *independent*: instead of a 6, we might as well have got a 4 or a 2.

This is a characteristically instable situation: an arbitrarily small difference of initial conditions will give completely different and independent results. This is the typical situation of a chaotic motion described in sec. 3.2, Fig. 3.4 corresponding fully to Fig. 3.3.

Even if we replace the human hand by a dice–throwing machine, the initial conditions will never be exactly the same, and the result is practically unpredictable. Theoretically its motion is determined by classical mechanics (if also the impact of the air molecules is considered a classical phenomenon), but prediction is hopeless. Laplace's demon, after having worried about the imprecise initial conditions, is then ad-

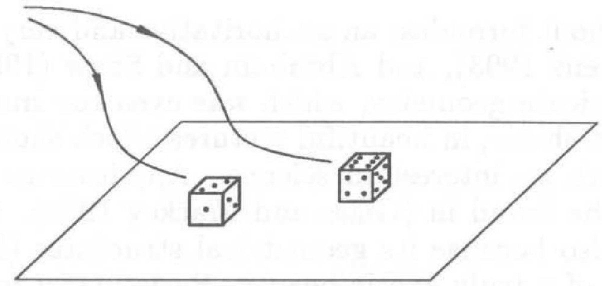


Figure 3.4: Throwing dice

ditionally bothered by Maxwell's demon (responsible for air molecules, cf. sec. 4.3).

The result of the fight between the two demons is a completely random distribution of the results of the die: face 1 is as probable as any other face. We may say that all faces have equal *probability*

$$p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = \frac{1}{6} \quad . \tag{3.21}$$

We see that Newton's laws, though theoretically applicable, are practically useless. *Exit* Newton, and *Symmetry* steps in and produces the result (3.21). More prosaically, determinism loses importance and symmetry takes over, producing *order out of chaos.*

(This is the reason why probability is treated here rather than in Chapter 2 where logically it would seem to be better placed.)

Had the die been *loaded*, then, of course, symmetry would have been destroyed and the probability of the various faces would be different. (We, of course, would never use such a dirty trick!)

Such assumptions of equal probability, based on symmetry, were used by Blaise Pascal (1623–1662) and contemporaries for a mathematical theory of games of chance. This was the foundation of the *mathematical theory of probability*. Laplace has perfected this symmetry–based theory.

Here the important concept of *symmetry* appears for the first time. A cube is symmetric because its six faces are geometrically equivalent: they can only be distinguished by marking them with dots, from one to six. If the faces were unmarked, then one could not distinguish a cube lying on face no. 2 from a cube lying on face no. 5. So much about *geometrical* equivalence or symmetry. A cube is also *physically* symmetric if it is made of a *homogeneous material*: this is what we mean by an unloaded die. A coin is symmetric if we disregard the inscriptions on the two sides: then we could not distinguishing a coin showing "head" from a coin showing "tail": we would in both cases see identical circles. We shall meet symmetry again; see secs. 3.6 and 4.2.

We have said that throwings of various faces were independent events. *Statistical independence* is a basic concept, though it by no means always holds. We shall, however, assume independence unless the contrary is asserted.

Throwing a 3 *or* a 5 has a probability which is the *sum*:

$$p(3 \vee 5) = p_3 + p_5 \quad . \tag{3.22}$$

Throwing a 3 *and* then a 5 is the *product*:

$$p(3 \wedge 5) = p_3 p_5 \quad . \tag{3.23}$$

These formulas do not presuppose equal probabilities (3.21), but they do presuppose independence.

Now we remember symbolic logic, eq. (2.9) on p. 28. The "logical sum" of two propositions was symbolized by "$\vee$", and the "logical product" by "$\wedge$". Now the *probability of a logical sum is the sum of probabilities* (3.22), and *the probability of a logical product is the product of probabilities* (3.23).

Probability 1 corresponds to certainty, and probability 0 to impossibility, and

$$0 \leq p \leq 1 \quad . \tag{3.24}$$

Obviously

$$p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1 \quad . \tag{3.25}$$

Thus the probabilities may be considered generalizations of or interpolations between the truth values 0 and 1, cf. (2.15) on p. 41. (Note the conflict of notations: in sec. 2.1, "p" stands for "proposition", here it denotes "probability". As a temporary compromise, we have in (2.15) symbolized probability by "P", but "p" is generally used in probability theory.)

This can also be nicely expressed in the language of set theory (Fig. 3.5). Throw a small particle at random in such a way that it lands on set A with probability $p(A)$ and on set B with probability $p(B)$. Both events may be considered independent if the two sets are disjunct. The union $A \cup B$ consists of both sets A and B taken together. Then

$$p(A \cup B) = p(A) + p(B) \tag{3.26}$$

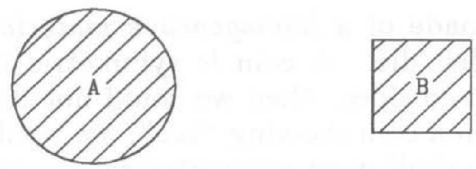


Figure 3.5: Two disjunct sets A and B

in analogy to (3.22); cf. (2.12) on p. 29. The corresponding relation

$$p(A \cap B) = p(A)p(B) \quad \text{(wrong!)} \tag{3.27}$$

unfortunately does not hold since the intersection $A \cap B = 0$ for disjunct sets. Here $p(A \cap B)$ would mean the probability that the particle lands *simultaneously* on A and B, which is clearly impossible, so that $A \cap B = 0$ implies $p(A \cap B) = 0$.

Remark on terminology. The terms "probabilistic", "statistic", "stochastic", and "random" have more or less the same meaning and are frequently used interchangeably.

Relative frequencies. Let us take an even simpler example, tossing a coin. For an ideally symmetric coin the probabilities p_1 of head and p_2 of tail are clearly equal:

$$p_1 = p_2 = \frac{1}{2} \quad . \tag{3.28}$$

If we throw the coin, say, a thousand times, there should be roughly 500 heads and 500 tails. In a real coin tossing experiment we may get, say, 484 heads and 516 tails. Thus the *relative frequencies* of heads and tails are

$$\begin{aligned} f_1 &= \frac{484}{1000} = 0.484 \quad , \\ f_2 &= \frac{516}{1000} = 0.516 \quad . \end{aligned} \tag{3.29}$$

If we throw 10000 times, we might get

$$\begin{aligned} f_1 &= \frac{5032}{10000} = 0.5032 \quad , \\ f_2 &= \frac{4968}{10000} = 0.4968 \quad , \end{aligned} \tag{3.30}$$

which is clearly closer to p_1 and p_2. It may be expected that, in some sense, for $n \to \infty$ throws

$$\begin{aligned} \lim_{n\to\infty} f_1 &= p_1 = 0.5 \quad , \\ \lim_{n\to\infty} f_2 &= p_2 = 0.5 \quad . \end{aligned} \tag{3.31}$$

In the so–called *frequency theory* of probability proposed by Richard von Mises since 1928, it was suggested to define probabilities empirically by such a limit

$$p \equiv \lim_{n\to\infty} f \quad . \tag{3.32}$$

This, however, meets with mathematical difficulties because the infinite limit does not obey one of the more common limit definitions of mathematics, and furthermore, it is not possible to perform infinitely many coin tosses or similar procedures.

It is mathematically simpler and more elegant to introduce the concept of probability *axiomatically.* This was done by A.N. Kolmogorov in 1933. Here the probabilities were introduced generally, without specifying their numerical values, but subject to axioms such as (3.22) and (3.23). Only later, approximate numerical values for them are found *a posteriori* as relative frequencies such as (3.30), unless they were not anyway given *a priori* by symmetry considerations (dice, coins).

The mathematical theory of probability has been developed to a high mathematical level, including random functions (stochastic processes) and Hilbert space techniques. Such techniques are, for instance, applied in geodesy to determine the irregular gravitational field of the Earth. This is called *least–squares collocation* and consists in an extension of least–squares adjustment (sec. 2.6) to infinite–dimensional Hilbert space. Only for curious specialists we mention as reference: H. Moritz: "*Advanced Physical Geodesy*", Wichmann, Karlsruhe, 2nd edition, 1990.

Interpretations of probability. In sec. 2.4 we have already briefly introduced *subjective probability*, expressing a degree of reasonable belief, or just a degree of incomplete knowledge or of ignorance. Of such character are the "probabilities of rain" given by American weather forecasts mentioned in sec. 2.4.

The classical Laplace interpretation is clearly intended to be *objective*. When I calculate my chance to gain in gambling to be 95% ($p = 0.95$), then I am not satisfied with this nice abstract result of mathematics: I expect to gain concrete money.

Are physical probabilities subjective concepts or objective features of nature? Consider statistical mechanics. In principle, presupposing the validity of classical mechanics, we could calculate the trajectories of all molecules without needing statistics. Statistics is needed because we cannot do this in practice. Hence we do introduce statistics just because of our inability or ignorance? This would indicate that our probabilities are more or less subjective.

On the other hand, statistical mechanics provides important "*emergent*" concepts such as temperature or entropy, and an elegant theory of thermodynamics has been developed on an axiomatic basis, without needing mechanics or statistics. It seems clear that temperature or entropy are objective "integral" properties of nature, and if they are derived by statistics, this statistics should be more or less "objective" as well. By the way, the derivation of thermodynamics from statistical mechanics is a beautiful example of the emergence of a *macro–law* from a *micro–law*. This is another example of *order out of chaos.*

A hundred years after Boltzmann, these questions are still being discussed. To be sure, the mathematical formalism and its results are completely unaffected by these "philosophical" discussions. Most working physicists could not care less whether their probabilities are subjective or objective. Weizsäcker (1985, p. 100) writes: "The concept of probability is one of the most striking examples for the 'epistemological paradox' that we can apply our basic concepts successfully without really understanding them."

Whether "deterministic chaos" on the basis of classical mechanics "really" introduces an objective probabilistic element into nature, is still an open problem under discussion. Every physicist, however, agrees that *quantum theory* does introduce objective probability into physics: quantum fluctuations form the basic substratum of our world.

Objective probability has been vigorously defended, also in quantum theory, by Sir Karl Popper (Miller 1985, sec. 15). He calls it *propensity* and interprets it in the sense of Aristotle's *potentialities* (possibilities) which are not all realized but are nevertheless *properties of nature*. In the progress of time, potentialities become actualities.

Summary. Probabilities have different interpretations, which are presumably all needed.

(A) *Probabilities of sets*. The current standard mathematical theory of probability, based on Kolmogorov's axiom system, is considering probability as a *measure of sets*. Any system of numbers which satisfies Kolmogorov's axioms is a possible system of probabilities.

Actual *estimations* of probabilities are done in two principal ways:

(1) By *symmetry considerations*. This is easy in the case of dice or coins, but may even be possible in complicated physical applications.

(2) By *relative frequency*. The toss of a coin regards the toss under consideration as one case out of an *ensemble* of 1000 or 10000 tosses, cf. equations (3.29) or (3.30) above. Similarly, in physics, our "real" physical system may be considered one out of a fictitious *ensemble* of possible "similar systems". This is the basis of Gibbs' approach to statistical mechanics (Lindsay and Margenau 1957, sec. 5.5).

Concerning the *physical reality* of probability or statistical considerations, there are two possibilities:

(a) Probability is only a function of *ensembles* of physical systems; probability considerations, such as in statistical mechanics, are only done statistically because a (deterministic) treatment is too complicated for us. Probability is a mathematical tool rather than a physical reality. Hence it may be regarded as *subjective*, at least to a certain extent.

(b) *Propensity*: this type of probability is a *physical property* of a single physical system, as *objective* as its mass, energy, or velocity.

(B) *Probability of propositions*. Mathematically they are very similar to probabilities (A), because the logical calculus of propositions is very similar to the logical theory of sets (sec. 2.1). "Subjective" or *subjectivist* probabilities of Carnap and others are of this type. (The forecast: "There is a 20% probability of rain for tomorrow" is a sentence, or in logical terms, a "proposition".)

Degrees of credibility. Not everything which we call probability must have a numerical value, or must be capable of being expressed numerically. If we say that all our knowledge is only probable, if we believe that the theory of relativity is very probably an outstanding theory, if I say that my train next day will probably run reasonably on schedule, it is difficult if not impossible to assign numerical values to the "probability" expressed by such statements. We instinctively act on beliefs with a high subjective degree of credibility as if they were absolutely true, and we disregard theoretical possibilities which are very small. When I go to work by car I know that I may have an accident. I take this into account in a reasonable way, by insuring my car, having my papers in order, and driving carefully. Having done this, I act as if this eventuality will not occur.

If I kept in mind all the possible events which theoretically might happen, but with a very low probability, then I would "probably" turn crazy or at least become a "professional worrier". This presumably is what Bishop Butler had in mind when he said that *probability is the guide of life* (Russell 1948, Part V, Chapter VI, p. 398).

In real life there is no absolute logical certainty, in the same way as there are no real mass "points", ideal straight lines or ideally exact measurements, cf. sec. 2.4.

Suggested additional reading. Probability, especially of the subjective type, frequently is treated together with induction (to be considered in sec. 3.9). Our standard reference (Lindsay and Margenau 1957) is slightly out of date on this topic but nevertheless worth reading. There are many excellent books on mathematical probability. An easy and delightful brief introduction by the most outstanding Russia specialists is (Gnedenko and Khinchin 1962). Geophysicists will not want to miss (Jeffreys 1961, 1973). A recent excellent discussion of all interpretations and their philosophical aspects is (Cohen 1989). Almost all aspects of probability in their historical development from Blaise Pascal to Niels Bohr are discussed with relatively little mathematics but with beautiful physical intuition in (Ruhla 1992). The remarks in Weizsäcker (1985, Chapter 3; 1992, Chapter 4) are brief but profound.

3.4 The theory of relativity

Henceforth space by itself, and time by itself,
are doomed to fade away into mere shadows,
and only a kind of union of the two
will preserve an independent reality.

Hermann Minkowski

Special relativity

Einstein's special theory of relativity deals with *inertial systems*. An inertial system according to Newton's theory is a system on which no force acts, so that equations (3.1) and (3.7) (p. 73–74) hold:

$$\ddot{\mathbf{x}} = 0 \quad ; \tag{3.33}$$

$$\mathbf{x} = \mathbf{a}t + \mathbf{b} \quad . \tag{3.34}$$

Eq. (3.33) says that there is no acceleration, and (3.34) says that the motion of an inertial system is *uniform*, that is with constant velocity along a straight line. An example is a spaceship in intergalactic space whose rockets have been shut off and which, of course, continues to move with constant velocity to reach the nearest galaxy.

A system which moves uniformly with respect to an inertial system, *is also an inertial system*. According to classical mechanics, the two are related by a Galilei transformation:

$$\begin{aligned} \bar{x} &= x - vt \quad , \\ \bar{y} &= y \quad , \\ \bar{z} &= z \quad , \\ \bar{t} &= t \quad . \end{aligned} \tag{3.35}$$

These equations can be directly found by inspecting Fig. 3.6.

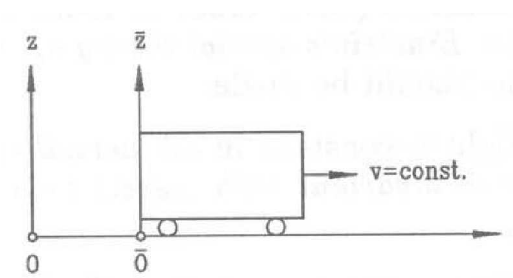


Figure 3.6: Two inertial systems in relative motion

Let a photon (particle of light) move with velocity c with respect to the car to which system $\bar{x}\bar{y}\bar{z}\bar{t}$ is attached, and let the direction of the photon coincide with the direction of velocity v. Then the velocity of the photon with respect to the original system $xyzt$ would be

$$c_0 = c + v > c \quad (?) \tag{3.36}$$

Now it has been extremely well confirmed experimentally that c is the maximum possible velocity. This is the *Michelson–Morley experiment*; cf. also secs. 3.8 and 3.9. It shows that the velocity of light in vacuum is constant and the same in every inertial system. Hence (3.36) cannot hold, there must be

$$c_0 = c \quad . \tag{3.37}$$

This, however, is incompatible with the formulas (3.35) for a Galilei transformation.

The correct transformation is

$$
\begin{aligned}
\bar{x} &= \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \quad , \\
\bar{y} &= y \quad , \\
\bar{z} &= z \quad , \\
\bar{t} &= \frac{t - \frac{v}{c^2}\,x}{\sqrt{1 - \frac{v^2}{c^2}}} \quad .
\end{aligned}
\tag{3.38}
$$

This is the *Lorentz transformation*. These formulas were first published by the Dutch physicist H.A. Lorentz in 1904 in connection with Maxwell's theory of electromagnetism and extended by Henri Poincaré in 1905. Its general validity for all physics, including mechanics, was recognized by Albert Einstein (1879–1955) in 1905. Thus classical mechanics was replaced by *Einstein's special theory of relativity*.

Several observations should be made:

(1) The velocity of light is constant in all inertial systems related by (3.38), that is, *c is invariant with respect to a Lorentz transformation*.

(2) For small velocities, $v \ll c$ or $v/c \ll 1$, the Lorentz transformation (3.38) practica lly coincides with the Galilei transformation (3.35). *Thus for "normal conditions", special relativity practically coincides with Newtonian mechanics*.

(3) *The Lorentz transformation is absolutely counterintuitive*: it is against "physical common sense". This is perhaps the first example of a counterintuitive physical theory (a second, even more counterintuitive, example is quantum theory), and has consequently been violently attacked, even by "conservative" physicists. This was at the beginning, around 1920–1930. Now, of course, special relativity is a fully established physical theory because

(4) Special relativity has been confirmed in all relevant cases without any exception. *It is perhaps the most accurately confirmed theory of physics*. This is particularly important for high velocities v almost $= c$, which occur with very fast–moving elementary particles in high–energy accelerators (CERN in Geneva, etc.).

Thus, a physicist denying special relativity would be taken even less seriously than a biologist denying the theory of evolution.

It is true that the mechanics of Galilei and Newton had also been considered counterintuitive by the defenders of Aristotelian mechanics, but a few very simple experiments were sufficient to recognize that Galilei and Newton were right.

Now, however, we are faced for the first time with the fact that *modern physics contradicts common sense.* This trend continues to the present day: physical theories are becoming more and more abstract and counterintuitive.

From a philosophical point of view this is extremely important: it shows that reality is much more complicated than we would have believed a hundred years ago. The positive outcome is that we get new models, however strange they look, which promote a much deeper understanding, not only of nature, but also of classical philosophical problems such as the nature of time and space, the structure of matter and even of mind. This particularly holds of quantum theory (sec. 3.5).

The line element. It may be shown that the *space–time line element*

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2 \tag{3.39}$$

is also invariant with respect to a Lorentz transformation (3.38). This is even more striking if we introduce a new "imaginary time coordinate"

$$x_4 = i\,c\,t \tag{3.40}$$

where $i = \sqrt{-1}$ is the "imaginary unit". Then, and on putting $x_1 = x$, $x_2 = y$, $x_3 = z$ as usual, we may write the line element in the form

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2 \tag{3.41}$$

which is nothing else than Pythagoras' theorem for dimension four, cf. (2.23) on p. 61! This was first recognized by the mathematician Hermann Minkowski (1864–1909) who wrote the beautiful statement used as a motto at the beginning of this section.

Thus space and time are welded together into a four–dimensional space, called space–time continuum or briefly, *space–time* (even written *spacetime*).

Note that the Pythagorean or Euclidean form (3.41) requires an imaginary time coordinate x_4. Much nonsense about this was and is written in science fiction: time is no more imaginary than space, time is no more unreal than space. What is true is that the imaginary transformation is simply a mathematical artifice to bring the line element into the Pythagorean form (3.41).

Beautiful as the form (3.41) is, however, it must be admitted that it may give a false impression that space and time are not different at all. An imaginary transformation, though easy mathematically, is a serious operation from the point of view of geometrical and physical reality. A *circle*

$$x^2 + y^2 = 1 \tag{3.42}$$

becomes, on replacing y by iy,

$$x^2 - y^2 = 1 \quad , \tag{3.43}$$

which is a *hyperbola* and thus a quite different geometrical figure. This simple example is relevant for relativity, as we shall see in sec. 3.7; cf. also (Moritz and Hofmann–Wellenhof 1993, pp. 181–195).

Anyway, it is now customary to use as time coordinate, not x_4, but

$$x_0 = ct \quad , \tag{3.44}$$

which is a simple real transformation of time scale, so that (3.39) becomes

$$ds^2 = -dx_0^2 + dx_1^2 + dx_2^2 + dx_3^2 \quad , \tag{3.45}$$

which is almost as simple as (3.41), but in which all quantities are *real*. It is called a *pseudo–Euclidean* form (mathematicians call it non–positive–definite).

General relativity

Surface theory. For introductory purposes, let us first replace four–dimensional space–time by a two–dimensional space, that is, a plane or a curved surface. The line element in the plane is, of course,

$$ds^2 = dx^2 + dy^2 \tag{3.46}$$

in Cartesian coordinates x and y. In polar coordinates r, ϕ with $x = r\cos\phi$, $y = r\sin\phi$ it becomes

$$ds^2 = dr^2 + r^2 d\phi^2 \quad . \tag{3.47}$$

In general curvilinear coordinates (u, v) it may be written

$$ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2 \quad . \tag{3.48}$$

This is quite natural: it must be quadratic, containing the squares du^2 and dv^2 and, in the general case, the mixed product $du\,dv$ as well. The quantities E, F, G, introduced by Carl Friedrich Gauss (1777–1855), are functions of u and v.

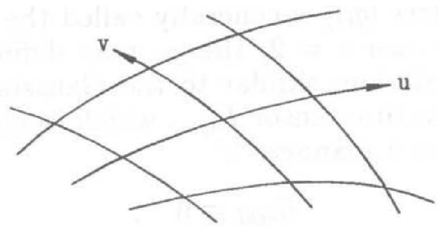


Figure 3.7: A curvilinear coordinate system

A curvilinear coordinate system is shown in Fig. 3.7. You may think of a grid of geographic coordinate lines on a map or on a globe.

In fact, (3.48) holds indifferently for curvilinear coordinates in the plane or on a curved surface. The distinction between a plane and a curved surface can be made only on the base of a certain highly nontrivial expression, the *Gaussian curvature* K, formed of E, F, G. If $K \equiv 0$, the surface is a plane, otherwise it is curved.

A more modern way is putting $E = g_{11}$, $F = g_{12} = g_{21}$, $G = g_{22}$ and $u = u_1$, $v = u_2$. Then (3.48) may be written

$$ds^2 = g_{11}du_1^2 + 2g_{12}du_1du_2 + g_{22}du_2^2 \tag{3.49}$$

or more shortly,

$$ds^2 = g_{ij}du_idu_j \quad . \tag{3.50}$$

Here the *summation convention* is used: a repeated index means summation over this index. In the present case, both i and j are summation indices running from 1 to 2.

Curved space–time. The form (3.50) is particularly important. It holds unchanged if the summation goes from 1 to 4 or from 0 to 3. Thus (3.50) represents the metric also for four–dimensional space time! If the coordinates u_i are now denoted by x_i ($i = 0, 1, 2, 3$), then

$$ds^2 = g_{ij}dx_idx_j \quad . \tag{3.51}$$

For instance, for the line element (3.45), the g_{ij} form the matrix

$$[g_{ij}] = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad . \tag{3.52}$$

Similarly to (3.49), however, the g_{ij} may also be functions of the coordinates x_i. The matrix $[g_{ij}]$ is generally called the *metric tensor*.

Similarly to the case $n = 2$, the g_{ij} may define a flat or a curved space–time. The criterion, similar to the Gaussian curvature, is now the Riemannian curvature tensor R_{ijkl}, which is again formed of the g_{ij} in a rather complicated manner. If

$$R_{ijkl} \equiv 0 \quad , \tag{3.53}$$

then we have "curvilinear coordinates" in the *flat* space–time of *special relativity*. These coordinates correspond to rotating or accelerating (non–inertial) coordinate systems. The above–mentioned space ship now no longer moves uniformly, but the rockets are working and changing speed, direction, or orientation in space of the space ship. A non–constant metric tensor $[g_{ij}]$ for which (3.53) holds, thus implies "*inertial forces*" in flat space–time, to which the crew of the space ship is exposed.

To give a more down–to–earth example, such "inertial forces" are well known from riding in an automobile which accelerates, brakes or sharply turns around a corner. To be sure, they are no "real" forces, but "only" due to having chosen a non–inertial reference frame, but tell this to a passenger who has been inconvenienced, endangered, hurt or even killed by such "unreal forces"!

The case of a non–zero Riemannian tensor is particularly interesting; we may write

$$R_{ijkl} = \text{gravitational field} \quad . \tag{3.54}$$

Curved space–time is equivalent to the existence of a "real" *gravitational field*, expressed (approximately) by Newton's law (3.11) on p. 74. This is the essential content of Einstein's *general theory of relativity*. Informally we may write (3.54) as

$$\textit{gravitation} = \textit{space–time curvature} \quad .$$

Newton's law has an equivalent in this theory. A certain quantity is formed from R_{ijkl}, called the *Einstein tensor* G_{ij}, and Einstein writes

$$G_{ij} = \kappa T_{ij} \quad , \tag{3.55}$$

where κ is a constant and T_{ij} is the *matter tensor*, more precisely the matter–stress–energy tensor. (The reader will have surmised in the meantime that "tensor" is any quantity that carries indices. He is basically right, at least for the present purpose. Let mathematicians protest against such a sloppy definition.)

Now comes a fact of true philosophical importance. The tensor g_{ij} represents geometry, R_{ijkl} represents curvature and hence also geometry, and so does the Einstein tensor G_{ij}, derived in a purely mathematical way from g_{ij}. *Hence the left–hand side of* (3.55) *represents geometry*. Since the tensor T_{ij} represents matter and κ is only a constant needed to get the dimensions right, (3.55) *expresses matter in terms of geometry*, more precisely in terms of space–time curvature.

What is more, this is the first logically impeccable definition of matter! The mass m in Newton's equations (3.1) and (3.11) was only an ill–defined constant.

Materialists will certainly not be satisfied with this highly abstract definition of matter by (3.55). This is not the type of matter we know, which we can kick, take into our hands, or throw, such as a stone or a ball. It is not, however, the only possible definition of matter as we shall see in sec. 5.1: there are worse definitions to come.

Preferred reference systems. In analytical geometry in the plane or in space, Cartesian coordinates are clearly preferable because of their simplicity. In special relativity, the same holds for *inertial systems*.

In general relativity, Einstein started from the *principle of covariance*: the equations of general relativity are covariant, that is, invariant with respect to general coordinate transformations. They have the same form in every coordinate system. In curved space–time, inertial systems are *theoretically* impossible; *practically*, systems can be defined which are very close to such inertial systems and thus form a kind of preferred systems. In Orwellian terminology, even in the perfect democracy of reference systems in general relativity, where all reference systems are equal, nevertheless some systems are more equal than others ...

By the way, this fact of the practical existence of inertial reference systems even in general relativity makes the transition to Newtonian mechanics possible. Thus general relativity is usually taken into account by applying, to the results of classical mechanics, some small relativistic corrections. They are on the order of 10^{-8} and can thus be measured by contemporary exact techniques. For instance, relativistic corrections are being routinely taken into account in the processing of precise geodetic satellite measurements, for instance in satellite laser ranging or also in precise uses of GPS, the well–known *Global Positioning System.*

Separability of gravitation and inertia. Besides the "real" force of gravitation, there are also "apparent" inertial forces due to a rotation or acceleration of the reference system (see above). For instance, people sitting in a car usually remain sitting comfortably because gravitation holds them to their seats, but they may start moving involuntarily if the car takes a sharp curve or brakes abruptly: seats are for gravitation and seat belts are against inertial forces.

From the time of Newton it has been recognized that gravitational and inertial forces act in very much the same way, so that they cannot even be separated. The reason is the equality between "inertial mass", the m in Newton's law of motion (3.1), and "gravitational mass", the m in Newton's law of gravitation (3.11). This equality has been verified experimentally to very high degree of accuracy, but theoretically it was a mystery until Einstein. In his general theory of relativity, Einstein cut the Gordian knot by boldly affirming that gravitational and inertial masses are identical because *gravitational and inertial forces are essentially the same phenomenon*. They arise *when the coordinate system in space–time is curvilinear* rather than inertial.

Again we have a nice Hegelian triad: *thesis* — inertial mass and gravitational mass are conceptually different; *antithesis*: but they are numerically equal; *synthesis*: this is because they are, after all, conceptually identical in the general theory of relativity.

At a single point, gravitational and inertial forces act inseparably together. In fact, ordinary *gravity* to which we are all subjected (and which is used to level geodetic theodolites etc.) is the resultant of gravitational attraction of the Earth and of the centrifugal force of the Earth's rotation. We do not notice the two different components of gravity, and no physical experiment can separate them.

If we wish, for some reason or other, to have only gravitation rather than gravity, we can calculate the centrifugal force, which is given by a very simple formula, and subtract it from measured gravity. Nothing is simpler than this.

If we wish to measure gravitational force in an airplane, however, matters are essentially more complicated: the inertial forces are so irregular that they cannot be directly determined or computed. The measurement of the gravitational force in an airplane is called *aerial gravimetry*.

When I got involved, around 1966, in the theory of aerial gravimetry, which attempts to separate gravitation and inertia, I casually talked to a physicist about this. He immediately said: "This is absolutely impossible: Einstein proved it". Then I looked up the current literature, and all books and papers confirmed this. The only exception was a book on general relativity by J.L. Synge, which had been on the market only for a few years. There I found eq. (3.54), and suddenly everything was clear.

The Riemannian curvature tensor R_{ijkl} does separate gravitation and inertia, cf. (3.53) and (3.54). It is formed by the metric tensor g_{ij} and its space–time derivatives of first and second order (non–mathematicians, don't get scared and forget it!). Such derivatives, however, need g_{ij} not only at one point, but also in a surrounding region, however small. Thus gravitation and inertia cannot be separated at one point only, but they can be separated in an arbitrarily small neighborhood of it! (The reason is that inertia has a much more regular space–time structure than gravitation.)

Components of the tensor R_{ijkl} can be measured, at least in principle, by instruments called *gradiometers*. Thus a combination of gravimeters and gradiometers provides, at least theoretically, a rigorous method for separating gravitation and inertia.

This is important, not only in aerial gravimetry, where gravitation is the "signal" and inertia is the "noise", but also in *inertial navigation* which is now used in almost every airplane: here inertia is the signal and gravitation is the noise. One person's signal may be another person's noise, as every musician knows from experience with his (her) neighbors.

Suggested reading. There are good nonmathematical presentations of the principles of relativity. My favorites are (Lanczos 1965) and (Will 1986). An "easy" but mathematical introduction is (Moritz and Hofmann–Wellenhof 1993). Remarkable as a very readable introduction to geometry in general is (Lanczos 1970). The treatment in (Lindsay and Margenau 1957) is excellent as usual. Philosophical problems of relativity are treated on a high level, also freely using physics and mathematics where necessary, in (Treder 1974).

3.5 Quantum theory

Natura [*non*] *facit saltus.*

Latin proverb

In the old small grocery stores it was possible to buy sugar, etc., in any reasonable amounts, e.g. 250 or 700 grams. In the modern supermarkets, sugar comes only in packs of 1 kg, say.

It seems that nature furnishes energy also only in fixed packages of a (of course, much smaller) constant size, called *quantum of energy*.

This discovery by Max Planck (1848–1947) in 1900 was applied to light by Albert Einstein in 1905. Light somehow behaves as if it consisted of identical particles, called *photons*. The same holds for the quanta of electricity, the *electrons*, for other elementary particles, and for the structure of atoms in general, as found by Ernest Rutherford (1871–1937) and Niels Bohr (1885–1962).

This was a heavy blow to classical physics, contradicting the old principle "*Natura non facit saltus*" (Nature makes no jumps).

The formalism of quantum mechanics

A (up to now) final theory of quantum phenomena was furnished in 1925–1926 by the almost simultaneous and largely independent work of several scientists of whom Werner Heisenberg (1901–1976), Erwin Schrödinger (1887–1961) and Paul Dirac (1902–1984) are best known. They did not develop three different theories, but three different aspects of the same theory.

The basic idea is the *correspondence principle*. A quantity in classical mechanics corresponds to a *linear operator in Hilbert space*.

About infinite–dimensional *Hilbert space* we have already spoken in sec. 2.6. A *linear operator* is the exact equivalence of a *matrix* in n–dimensional Euclidean space; it may be considered an infinite–dimensional matrix. (If you are scared, just skip the mathematics: we shall be finished with it in a few minutes.)

A (symmetric square) matrix can always be brought into a diagonal form:

$$L = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \lambda_n \end{bmatrix} \quad . \tag{3.56}$$

The λ_i are called *eigenvalues*. For linear operators, we simply have $n \to \infty$.

Physically it is basic that the *eigenvalues of a certain operator constitute the possible measured values of the physical quantity represented by our operator*.

A link to Hamiltonian classical mechanics (sec. 3.1) is the fact that any *canonically conjugate variables* q_k and p_k (generalized coordinates and momenta) are, in quantum theory, replaced by operators Q_k and

P_k satisfying the *commutation relations*

$$P_kQ_k - Q_kP_k = \frac{h}{2\pi i} \quad . \tag{3.57}$$

For the mathematically minded reader we explain the symbols: $i = \sqrt{-1}$ is the imaginary unit and h is a fundamental physical constant called *Planck's constant*. The operators P and Q, like matrices, are not in general commutative (such as $ab = ba$ for ordinary numbers). For $h = 0$, (3.57) reduces to commutative relations, $P_kQ_k = Q_kP_k$. In this case, quantum mechanics reduces to classical mechanics.

Operators act on functions just as matrices act on vectors. The equation

$$\mathbf{L}\,\mathbf{e} = \lambda\mathbf{e} \tag{3.58}$$

defines the *eigenvectors* $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ of the matrix $\mathbf{L}$, and the corresponding numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ are, by definition, the eigenvalues occurring in the diagonal of the matrix (3.56).

Let now L denote a linear operator. The precise analogue of (3.58) is

$$L\psi = \lambda\psi \quad . \tag{3.59}$$

The function ψ is called *state function*, and there is usually an infinite set of such ψ–functions: $\psi_1, \psi_2, \psi_3, \ldots$ (called *eigenstates* of operator L). The corresponding numbers $\lambda_1, \lambda_2, \lambda_3, \ldots$, the eigenvalues of L, are the possible outcomes of measuring the physical quantity represented by the operator L as mentioned above.

The mathematics, strange as it looks at first glance, is in reality quite simple. As we have already remarked, *linear operators* (according to Schrödinger) can be represented as *infinite matrices* (according to Heisenberg), and Schrödinger's *state functions* as Heisenberg's *infinite state vectors*. This analogy is now well known to any student of theoretical physics, to many applied mathematicians (Fourier transform!), and also to theoretical geodesists, cf. (Moritz and Hofmann–Wellenhof 1993, Chapter 6). Matters are far less simple with

Problems of interpretation

State functions. Such functions ψ were introduced by the Austrian Erwin Schrödinger; therefore the symbol ψ quite appropriately decorates the Austrian 1000 Schilling banknote showing his portrait. This, however, is not the most important property of the ψ–function; much more important is the fact that it is difficult to find two prominent physicists or philosophers of science who agree on its meaning.

Schrödinger himself was dissatisfied with all current interpretations throughout his life. This is another striking example of Weizsäcker's "epistemological paradox" mentioned in sec. 3.3: we can perfectly well work with concepts whose real significance we do not fully understand.

Suppose we measure a quantity L and get the value λ_3. Then we know that the physical state of the system is described by the state function $\psi = \psi_3$. It will evolve with time according to an equation found by Schrödinger, giving $\psi(t)$. Let us further assume that $\psi(t)$ is "normed" to correspond to a *unit vector*.

Now we perform the measurement of another physical quantity, represented by the operator M with eigenvalues μ_k and normed eigenfunctions $\xi_k(t)$. Then *one* of the eigenvalues μ_k *must* be observed, but we don't know beforehand which. But now comes the principal point: we know the *probability* that a certain μ_k is measured. This probability is precisely

$$p_k = <\psi, \xi_k> \quad , \qquad (3.60)$$

where the symbol $<>$ denotes the inner product of two state functions. If you know what the inner products of two vectors is, think of ψ and ξ_k as vectors, and you have it; otherwise let it be just another of those crazy mathematical terms and forget it. Thus we do not know whether, say, ξ_2 or ξ_9 is measured, but we know the probabilities for ξ_2, namely

$$p_2 = <\psi, \xi_2> \quad ,$$

or for ξ_9, namely

$$p_9 = <\psi, \xi_9> \quad . \qquad (3.61)$$

If tonight we listen to the weather forecast for tomorrow, it might be: $p_2 = 80\%$ probability for rain and hence $p_9 = 20\%$ for absence of rain. (We are purposely using p_2 and p_9, which is better for the quantum analogy, rather than p_1 and p_2 as we should normally do.) Next morning we look out of the window: sun is shining brightly. Thus the event corresponding to the probability p_9 has been realized (maybe, St. Peter wanted to show the meteorologists who was who).

To return to our quantum measurement. We perform the measurement of M and the outcome is μ_9, and hence the new state function is ξ_9, for the very same reason that after measuring L with result λ_3 the state function was ψ_3.

Thus, as the result of our measurement, the state function, having been ψ_3 before the M–measurement, has suddenly become ξ_9 afterwards.

It thus turns out that measurement changes quantum states discontinuously, from ψ_3 to ξ_9. *Natura facit saltus*!

Popularly speaking, a measurement disturbs a quantum state, and it disturbs it unpredictably. From a physical point of view this is quite plausible. Observing an electron by a microscope, say (assuming this possible), means interaction of the electron with light, that is, we have a collision between an electron and a photon. The outcome is as unpredictable as with the collision of two cars, which is also subject to probabilities (on the basis of which insurance companies calculate their premiums).

This is expressed in the famous *Heisenberg uncertainty relation*. They are a direct consequence of the commutation relations (3.57). Any two canonically conjugate variables p and q cannot be measured accurately *at the same time*. There are unavoidable measuring errors Δp and Δq related by the uncertainty relation

$$\Delta p \Delta q \doteq h \tag{3.62}$$

where h is again Planck's constant. If q denotes a coordinate x, and p the corresponding momentum $p = m\,v$ (mass times velocity), Heisenberg's uncertainty relation states that, if we measure the position $x = q$ accurately ($\Delta q = 0$), then the momentum p becomes indeterminate ($\Delta p \doteq h/\Delta q \to \infty$) and vice versa. In fact, the light quantum used in the microscope to measure the position $x = q$ of the electron collides with it and thus unpredictably disturbs its momentum p (size and direction of its velocity). The measurement of coordinate q and the measurement of momentum p are *complementary events*. In a related way we say that the wave picture and the particle picture mentioned below are complementary. *Complementarity*, introduced by Niels Bohr, is closely related to *dialectics* (sec. 2.5).

But let us return to the discontinuous change from the state $\psi = \psi_3$ to ξ_9. The difference with the car collision is that the car accident is subject to the laws of mechanics and can therefore be precisely and "deterministically" described by them (in principle!). In the case of the electron–photon collision, there are no underlying deterministic laws; probability is essential and cannot be reduced to some "deeper" deterministic laws.

Thus we have two ways in which a state can change: a continuous "deterministic" evolution according to an equation called "Schrödinger equation", denoted by **U** (for "*U*nitary transformation") by Penrose (1989, p. 250), and a discontinuous irreducibly random change by the observational process, called **R** by Penrose ("*R*eduction of the wave packet").

The term, wave packet, refers to the fact that, in a way, ψ represents "matter waves". Electrons, besides being particles, in other circumstances behave as waves (the "matter waves" of de Broglie and Schrödinger), in much the same way as light, usually appearing as waves, in certain respects also has particle character; there correspond:

light waves and photons,
matter waves and electrons.

The matter wave picture also beautifully illustrates the Bohr–Rutherford model of the hydrogen atom consisting of an electron orbiting around the atomic nucleus (Fig. 3.8). A stable orbit must form a "stationary wave" containing an integer number of wave lengths (orbit 1). Orbit 2 does not satisfy this condition and hence is not possible: the corresponding matter wave is discontinuous.

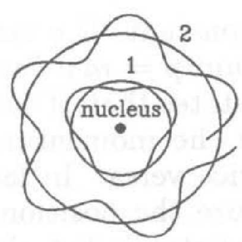


Figure 3.8: Orbit 1 is stationary (closed) and hence possible, orbit 2 is not closed and thus is impossible

Stable (stationary) orbits thus form a *discrete* set: only a finite number of orbits correspond to stationary waves. Hence *discreteness* arises out of the *continuity* of waves: another beautiful example of dialectics as discussed in sec. 2.5!

If the light is so weak that there is only one photon, then the light wave may be considered to describe the probability of incidence of a photon, say in a photon detector. In the same way, a matter wave as represented by a function ψ may be regarded as a *probability wave* describing the incidence of an electron.

The role of the observer. We have seen that the quantum observation has thrown the quantum state ψ_3 into the new state ξ_9 (Penrose's transformation **R**). To return to our meteorological example, has our looking out of the window changed the weather suddenly to sunshine? In meteorology this opinion would be considered insane, in quantum theory it is the generally accepted view.

What is causing "the reduction of the wave packet **R**"? Any answer given so far seems to raise problems, especially since the observer's *mind* seems almost always to be involved.

Schrödinger's cat. Schrödinger gave a famous striking example, a "thought experiment". Consider a cat sitting in a black box, without being capable of being observed from the outside. In the box there is a radioactive atom which is known to decay at an unknown random instant. Let it be known that the atom, with equal probability, decays or not during the couple of hours the cat is in the box (well equipped with food, of course). If the atom decays, it triggers a device which smashes a phial containing cyanide, killing the cat. At the end of the experiment the observer opens the box and finds the cat dead or alive. According to the usual quantum interpretation, only this observation is determining whether the cat is alive or dead; before opening the box, the cat cannot be said to be either alive or dead!

By means of this paradox, Schrödinger tried to express his unsatisfaction with current interpretations of quantum mechanics.

Ingenious solutions of the cat paradox have been offered, none of them entirely convincing: otherwise it would not be discussed even at the present time (it was published in 1935!). Alluding to the standard "Copenhagen interpretation" of quantum mechanics, largely due to Niels Bohr, Schrödinger ironically remarked: "I think I must accuse Bohr — though in actual fact he is one of the kindest persons I ever came to know — of unnecessary cruelty for his proposing to kill his victim by observation."

The Copenhagen interpretation. This interpretation, developed mainly by Bohr together with Heisenberg, has been adopted by the majority of physicists, more or less consciously. It has already been used implicitly in much what has been said above. The wave function ψ incorporates an *objective* and a *subjective* aspect, which are related to Penrose's transformations **U** (a "unitary transformation" or a rotation in Hilbert space: objective) and **R** ("reduction of the wave packet" by the observation: subjective). The observer as a subjective element enters essentially, at least because he designs the extremely complicated apparatus needed, but many physicists (e.g. E. Wigner) also ascribe an essential role to the *mind* of the observer to register and thus fix the outcome of the experiment: this is precisely what Schrödinger is alluding to in the "accusation" just quoted that Bohr "proposed to kill his victim" (Schrödinger's cat) by observation. A classical brief but ideally competent and readable presentation of the Copenhagen interpretation is found in (Heisenberg 1958, Chapter III).

One of the difficulties with the role of the observer is the question how quantum theory applied to nature before the advent of man. Have dinosaurs also been subjected to the laws of quantum mechanics? The answer of the Copenhagen interpretation is that the question is meaningless: quantum theory describes, *not nature as such, but our interaction with nature* by means of experiments: dinosaurs probably did not perform quantum mechanical experiments. As Weizsäcker put it: nature was before man, but man was before quantum mechanics. But then the question arises again: which law took the place of quantum mechanics before quantum phenomena came to the attention of man around 1900?

It seems that such objections do not arise, at least not so conspicuously, in the various "objective" interpretations which we shall discuss now.

Every interpretation proposed so far contains some strange feature. The strangest probably is the

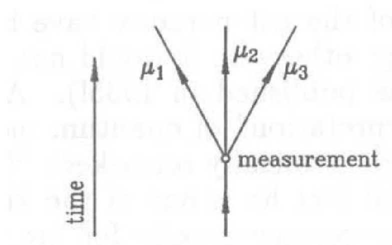


Figure 3.9: Illustrating Everett's interpretation with $n = 3$: the universe branches out into three equally real worlds

Many–world interpretation. This interpretation, proposed by H. Everett in 1957, tries to avoid the arbitrariness involved in the realization of just one possibility: in the above example, of ξ_9. Why not ξ_3 or ξ_{15}? Everett says that in fact *all* possibilities are realized simultaneously. Hence with n (or infinitely many) possible outcomes of the measurement (number of eigenvalues), the world forms n (or $n \to \infty$) branches (Fig. 3.9). At each measurement, the branching out is repeated (Fig. 3.10). Thus the world containing Schrödinger's cat branches into a world in which the cat is alive and an almost identical world in which the cat is dead.

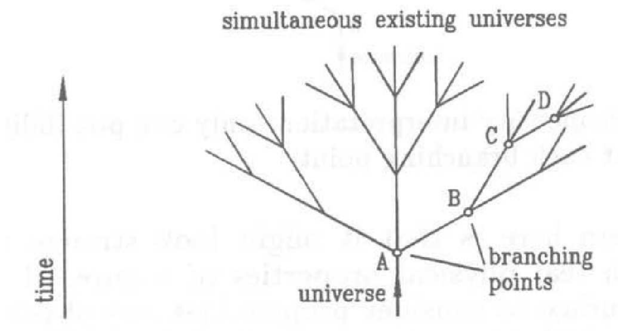


Figure 3.10: Repeated branches, all being equally real

It has been shown that this interpretation is perfectly logical and self–consistent. It is also possible from a physical point of view (conservation of energy). According to Bohr's dictum quoted as the motto of sec. 6.6, Everett's theory may well be "crazy enough to be true".

Whitehead once said that it was more important that a proposition were interesting than that it were true. In fact, Everett's many–world interpretation is taken seriously by a small but nevertheless surprising number of most outstanding physicists such as Feynman, Gell–Mann (1994, Chapter 11), Hawking, and Weinberg. It avoids the "reduction of the wave packet" of the standard interpretation by replacing it with a perfectly "objective" interpretation. Thinking, however, that the incessant branching of the "world lines" creates an infinite number of parallel worlds of which many are similar except for some detail and some are quite different, makes many people feel dizzy and ask for an application of "Occam's razor" to cut off all branches except one. This is the purpose of the

Propensity interpretation advanced independently by the philosopher Karl Popper (Miller 1985, sec. 15) and the physicist Robert Havemann (1964, pp. 96–103). Here the probabilities are "propensities", objectively existing tendencies. The difference with respect to the many–world interpretation is that only *one* of the n possibilities is realized, cf. Figs. 3.10 and 3.11. The probability of the branch chosen is, of course, given by (3.60), now interpreted as a propensity.

The problem here is that it might look strange to consider the propensities as real physical properties of nature. It may, however, be only a prejudice to consider propensities less physical than, for instance, forces. Both concepts are equally "physical" or equally "meta-

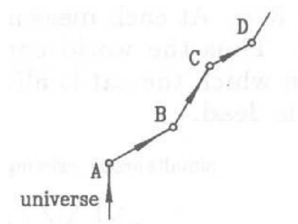


Figure 3.11: Propensity interpretation: only one possibility is realized at each branching point

physical", as Popper says. On closer look, also concepts like mass or force lose their intuitive meaning acquired from every–day experience with manual labor.

As we have said in sec. 3.3, the philosophical foundation of Popper's view is the Aristotelian distinction between potentiality and actuality (reality), putting "the propensities as potentialities into things". This is clearly related to ideas of R. Bošković, cf. sec. 5.2, p. 210.

The physicist Havemann arrives at almost the same interpretation starting from the Hegelian dialectics of potentiality and reality. By the way, the booklet (Havemann 1964) appears to me to be by far the best application of Hegelian and materialistic dialectics to contemporary physics.

The ensemble interpretation. For dialectic materialists in the former Soviet Union, the propensity interpretation seems to have been not "materialistic" enough. There is, however, no logical difficulty in replacing the propensity interpretation of probability by a frequency interpretation (sec. 3.3), ascribing probability not to one system, as Popper did, but considering an *ensemble* of similar systems. The little but beautifully written book (Blokhintsev 1968) appeals to the physicist but is also very instructive and highly pertinent to our philosophical discussion.

The holistic interpretation. As recognized already by Schrödinger, quantum mechanics has a strong holistic flavor: "the whole — Greek *holos* — is more than the sum of its parts" (cf. sec. 2.1): a quantum system cannot simply be decomposed into a sum of independent subsystems. Holism comes natural to Eastern — Indian, Chinese, Japanese — philosophy, as particularly well (Rabbi type 1) described in (Capra 1975) and (Moser 1989). The most profound presentation (Rabbi type 2) is (Bohm 1980). (For "Rabbi types" see the Preface.)

According to Bohm, the real world is much more complex than our three–dimensional space or four–dimensional space–time. Loosely speaking, the "many worlds" of H. Everett are *enfolded* into a higher–dimensional universe from which our space–time world is only a certain projection down to four dimensions. Quantum effects can propagate faster than light (without, of course, carrying information for which the light velocity c is the limit after all). A very rough model, which does not pretend in any way to be realistic, is shown in Fig. 3.12. Besides the normal space–time route with maximum velocity c there is a shortcut in the form of a "jump across hyperspace" AbB, to follow the terminology of the well–known "*Foundation Trilogy*" by the science–fiction author Isaac Asimov.

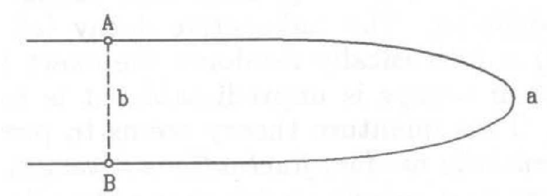


Figure 3.12: A "fold" in the universe. Effects can propagate along the "normal" space–time route AaB, but also "cut across" the fold following route AbB

The shortcut AbB demonstrates the "*nonlocality*" of quantum mechanics, which exists in *any* interpretation of quantum mechanics but comes out particularly well in the holistic interpretation. It says that the universe is essentially more complex than the four–dimensional space–time of Albert Einstein. Einstein had many discussions with Niels Bohr, in the course of which he invented, as a thought experiment, the *EPR* (Einstein–Podolsky–Rosen) *paradox*, no less famous than Schrödinger's cat (remember, it was "killed by observation"). The general quantum–mechanical fact underlying non–locality was expressed 1965 by John Bell in the *Bell inequality*, which was empirically tested by Aspect and others in 1982 and several times since.

Bohm's interpretation with its hidden enfolded reality reminds of Kant's "Ding an sich" (thing in itself) which remains forever unknown (sec. 5.4).

Concluding remarks and further reading. Quantum theory and relativity, particularly Einstein's special theory, are as well tested as classical mechanics to which they reduce for ordinary macroscopic systems ($h \to 0$) and small velocities ($v/c \to 0$). So far, within the limits of their validity, none of these theories have ever been contradicted by experiment: Popper's "falsification" (sec. 3.9) has still to come for them.

For the working physicist, the results hold independently of the philosophical interpretation. The philosophical interpretation of quantum theory is much more difficult and controversial (and therefore even more interesting) than the interpretation of relativity. The interpretations of quantum mechanics are closely related to the various interpretations of probability. All interpretations described above seem to be logically impeccable; their preference may be a matter of taste. (I like them all.)

The basically probabilistic character of quantum theory holds, of course, independently of interpretation. The probabilities (3.60) of transition from state ψ to state ξ_k exist and cannot be replaced by something more definite. The radioactive decay (cf. the example of Schrödinger's cat) is intrinsically random: the exact instant in which the radioactive atom decays is unpredictable; it is subject only to a probabilistic law. Thus quantum theory seems to provide *a universal background of essentially random fluctuations* ("vacuum fluctuations").

Quantum theory particularly clearly shows the relevance of modern physics to philosophical questions such as the nature of matter and mind, reality, potentiality, and probability.

It shows, however, also the relevance of philosophy to physics: we have met with ideas of Aristotle, Kant, Hegel, dialectic materialism, and Eastern philosophy; and the overarching influence of Plato will become evident later.

I cannot resist the temptation to call quantum theory a dialectic synthesis of extreme (mathematical) simplicity and extreme (philosophical) complexity.

The practical relevance of quantum theory is remarkable. Among many other achievements, it allows to *reduce chemistry to physics*. Quantum mechanics explains the structure of the atoms of chemical elements, the formation of molecules from such atoms, i.e., the nature of chemical bonds, etc.

Besides the books already quoted, we mention some more general works. It is recommended to first read (Davies and Brown 1986), a wonderfully readable (Rabbi type 1) introduction and overview of the major interpretations with interviews of the leading proponents. (Gribbin 1984) comes next in readability. Bohr's (1958, 1963) books are basic but rather difficult (Rabbi type 2). Interpretation problems of quantum theory with emphasis on the Copenhagen interpretation are dealt with extensively in the books by Heisenberg and Weizsäcker which are much more readable. Contemporary reviews are (Lockwood 1989) and (Stapp 1993), both a "must" for those who wish to penetrate more deeply into the mysteries of quantum theory (Rabbi type 1 to 2). (Penrose 1989) is excellent but rather difficult (almost type 2). (Schrödinger 1958) is a classic, popular in the best sense.

3.6 Elementary particles

Three quarks for Muster Mark!
Sure he hasn't got much of a bark
And sure any he has it's all beside the mark.

James Joyce, Finnegans Wake

History and description

The first elementary particles were the *electron* (the quantum of negative electricity) discovered by J.J. Thomson in 1897, the *photon* (quantum of light) found by Einstein in 1905, and the *proton* (the nucleus of the hydrogen atom) which appeared first implicitly in the Bohr–Rutherford model of the hydrogen atom around 1910–1913. The hydrogen atom was considered to have a nucleus consisting of a positively charged proton, around which the negative electron is orbiting in various "permissible" orbits. The proton was discovered explicitly by Rutherford in 1919. Curiously enough, the nucleus of the helium atom consisting of two protons, had been isolated already around the turn of the century and was known by the name of α–particle (β–particles were electrons, and γ–particles are photons).

The mass of the positively charged proton was found to be about 1836 times as great as that of the electron. In 1932, the British physicist James Chadwick discovered the *neutron*, having a very similar mass (1840 electron masses) but being electrically neutral, hence its name.

All atoms were found to consist of protons, neutrons, and electrons. Since all atoms could be reduced to these three particles, it was believed that these three were the only ultimate constituents of nature, the only elementary particles. They were described by the recently discovered quantum theory, and together with relativity the physical explanation of the universe appeared to be essentially finished. Physics had become a routine job elaborating the details of atomic and molecular structure by known experimental and theoretical methods. The end of physics appeared to be in sight: its job had almost been done.

Such stages of "near completion" of a science in general and physics in particular are well known from the history of science. They frequently signal the transition to a new "paradigm" in the sense of Kuhn (sec. 3.10). The mechanist world picture of Newton–Laplace seemed to explain almost everything in physics, even heat as statistical mechanics and electromagnetism as the mechanical oscillations of an elastic "ether". This mechanist picture reigned supreme: the discovery of relativistic and quantum phenomena came entirely unexpected.

When Stephen Hawking assumed the famous Lucasian chair in Cambridge (previously held by Newton and Dirac) in 1979, the title of his inaugural lecture was: "Is the End in Sight for Theoretical Physics?" The new unified theories of supergravity and superstrings seemed to indicate that we might be close to a final physical theory. Now we seem to be farther away from this goal than ever before, cf. sec. 6.6.

But let us return to the situation with quantum theory and atomic and particle physics around 1930. Photons and the constituent particles of atoms: electrons, protons, and neutrons were all known.

But in the early thirties, new elementary particles were discovered: the elusive *neutrino* which practically does not interact with matter, so that even the Earth is "transparent" to it, and it passes straight through; and the *positron*, the positive analogue to the electron. Thus in 1933, already 7 particles (proton, neutron, electron, positron, 2 kinds of neutrino, and photon) were known.

The *photon* is the quantum of light, that is, of electromagnetic waves. The *electromagnetic force* binds electrons to the atomic nucleus. In 1935, the Japanese physicist H. Yukawa suggested that the *nuclear force* which holds protons and neutrons together in the atomic nucleus, also corresponds to a particle, called *meson*.

By 1947, some 14 elementary particles were known. Besides the 7 mentioned above, there were the "antiparticles" *antiproton* and *antineutron*, and 5 kinds of *mesons*.

Some of them were found in cosmic radiation, but the decisive experimental event was the construction of more and more powerful particle accelerators such as cyclotrons and linear accelerators. In these machines, the particles (protons, electrons, charged heavier atomic nuclei) are accelerated to extremely high velocities and correspondingly high energies; they are made to collide with each other, and new particles are produced in this high–energy collision. Famous in Europe is CERN (European Council for Nuclear Research) in Geneva, Switzerland.

Now the number of known elementary particles started to increase exponentially. Instead of the atomic–particle triad around 1930 (proton, neutron, electron), we now know hundreds of elementary particles. Scientists have been expelled from the original particle paradise to find themselves in a big particle zoo full of strange and extremely expensive animals.

Even now, most properties of the chemical elements, of their molecules and of the atoms of which they are made up, can be explained by the triad: proton, neutron, and electron. The other particles occur naturally in cosmic radiation; most of them, however, are "manufactured" artificially in the big accelerators. Nevertheless they are also essential for at least two reasons:

(1) They may occur in the process of formation of stars.

(2) They are necessary to set up general unified theories of physics and to test them experimentally.

New particles now are hardly discovered accidentally; they are predicted by theory and then systematically searched for by experiment.

Antimatter. In 1928 Dirac found his famous equation for the electron, which was the first successful unification of quantum theory with *special* relativity (a convincing unification of quantum theory with general relativity is still an open problem). It was soon discovered that Dirac's equation admitted a second solution, identical to the electron except for its electric charge. This is the *antielectron*, called positive electron or briefly *positron*.

Soon it was found that each elementary particle has its anti–counterpart, called *antiparticle*. There is a negative *antiproton*, even an *antineutron* (also electrically neutral, but electric charge is not the only criterion). Only the antiparticle of the photon is the photon itself.

When a particle meets its antiparticle, they mutually annihilate each other, generating a large amount of energy, according to the famous equation of Einstein: $E = mc^2$, energy equals mass times the square of the light velocity. Remote galaxies which are apparently in violent explosion may be due to a galaxy consisting of ordinary matter colliding with a galaxy consisting of antimatter. At any rate, this happens in science–fiction, in which antimatter is very popular. Fortunately, in our world antiparticles usually do not occur naturally; this unsymmetry or "*symmetry breaking*" between particles and antiparticles is rather mysterious.

Spin. An elementary particle possesses an intrinsic angular momentum, called *spin*, which is not due to actual rotation in space but is an abstract quantum phenomenon. The spin number of a particle may be one of the values

$$s = 0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ 2,\ \frac{5}{2},\ \ldots \tag{3.63}$$

and their negative values. In fact, the angular momentum corresponding to spin s is

$$s\frac{h}{2\pi} \tag{3.64}$$

which shows spin to be a typical quantum phenomenon (it becomes zero if $h \to 0$, in the transition from quantum mechanics to classical mechanics). Mesons have spin 0, electrons, protons, and neutrons have spin 1/2, photons have spin 1, and gravitons (the hypothetical particles corresponding to the gravitational force) would have spin 2.

Elementary particles can be classified by their spin. Particles with integer spin are called *bosons*, the others *fermions*. A different division is between leptons and hadrons.

Leptons (Greek: particles of light weight). There are 6 leptons which all have spin 1/2: the *electron*, and two similar but much heavier and extremely shortlived particles (*muon* and *tauon*) and three *neutrinos* corresponding to the first three negative particles (we are here disregarding the corresponding antiparticles).

Hadrons (Greek: strong particles). They consist of three groups:

(1) *Mesons*: particles of medium mass, especially the pi–meson (muons and tauons are sometimes also called mesons).

(2) *Nucleons*: proton and neutron.

(3) *Hyperons*: they are heavier than nucleons and extremely shortlived (so–called Λ, Σ, Ξ, and Ω particles).

Nucleons and hyperons together are called *baryons* (Greek: heavy particles).

Hadrons consist of smaller particles called *quarks* (whimsically called after the quotation by James Joyce which figures as the motto of the present section). There are some 18 different types of quarks with deliberately nonsensical names, of different masses and different electric charges and other properties.

Mesons are formed by the combination of *two* quarks, and *baryons* (nucleons and hyperons) are made up of *three* quarks. (Leptons seem to be truly elementary, not made up of any smaller particles.)

Orders of magnitude. A typical size of a molecule is 1 nm (1 nanometer = 10^{-9} m), of an atom 10^{-1} nm, of an atomic nucleus 10^{-5} nm, of a proton 10^{-6} nm, and of a quark less than 10^{-9} nm. Electrons and quarks are practically pointlike. (Remember that the wavelength of visible light is about 500 nm (sec. 1.3).) The size of an atom is about 10000 nuclear radii, and the size of a proton is at least 1000 times larger than the size of a quark. (This is even greater than the ratio between the orbit of the Moon and the size of the Earth, and between the orbit of the Earth and the size of the Sun!) Thus both an atom and a proton consist essentially of empty space in which, however, extremely strong forces act.

The four forces. The basic forces occurring in nature are:

– *gravitation*,

– *electromagnetism*,

– the *weak force* and,

– the *strong force*.

Gravitation and electromagnetism are sufficiently known to us. The *weak force* occurs in the interaction between *leptons* or more popularly, in radioactive phenomena. The *strong force* is responsible for the interactions between *hadrons*; it holds the nucleus together.

Exchange particles. The electromagnetic interaction between nucleus and electrons may be regarded as an exchange of *photons*, as we have already remarked. Similarly, a quantum theory of gravitation has been seen to lead to *gravitons*. Yukawa looked for the *meson* as exchange particle for the *strong force* between nucleons. The *weak force* (between leptons) similarly corresponds to what has been called *W and Z particles* which are extremely heavy: 80–90 proton masses. Finally, *quarks* are held together by *gluons* (English "glue" with a Greek touch).

Summary. This all sounds quite complicated, so what should you really know? *Most particles are fermions* of spin 1/2: electrons, protons, neutrons (generally all leptons and baryons), as well as quarks. *All exchange particles are bosons*:

electromagnetism	photon	spin 1
gravitation	graviton	spin 2
weak force	W^-, W^+, Z	spin 1
strong force	gluons	spin 1
	[mesons	spin 0]

(Yukawa's mesons (pi–mesons or pions) have been considered responsible for the strong force between neighboring nucleons (protons and neutrons), whereas gluons act between quarks. Now Yukawa's theory may appear superseded by the gluon theory of the strong force, but its pioneering role is universally recognized.)

Symmetry

The only way to bring order into this chaotic zoo of particles is by means of various symmetries, which admittedly arise in a rather abstract and not very intuitive way.

Symmetries are mathematically expressed by *groups of transformations*, acting on an object $\mathbf{x}$ (which may be some abstract vector). The simplest group consists only of the *identity transformation* $\mathbf{I}$, leaving the object unchanged

$$\mathbf{I}\,\mathbf{x} = \mathbf{x} \quad . \tag{3.65}$$

This identity group is responsible for the identity of all particles of a certain kind: all electrons are identical, so are all protons, etc. This fact is by no means trivial, but it corresponds to physical intuition.

Rotation in two–dimensional real space, in the plane, curiously enough, is related to *electromagnetism*, via its gauge theory, see below. (This relation is absolutely counterintuitive!)

Rotation in three dimensions plays a great role in the study of rotationally–symmetric *atoms*. This group is denoted by $O(3)$. (This is abstract but reasonably intuitive.)

Already ordinary quantum theory is forced to operate with (auxiliary) complex space, in which the coordinates are complex numbers of form $a+bi$, $i = \sqrt{-1}$ (sec. 3.5). A *unitary transformation* is in complex space what rotation is in real space.

The group of unitary transformations in *two–dimensional complex space* is denoted by $SU(2)$ (actually $U(2)$ but we shall here always disregard mathematical niceties; $SU(2)$ means "Special $U(2)$ "). Curiously enough, $SU(2)$ is mathematically equivalent to the *real rotation group in three dimensions*:

$$SU(2) \hat{=} O(3) \quad . \tag{3.66}$$

The group $SU(2)$ is responsible for *spin* and related properties. We mention in passing that, in some sense, $SU(2)$ is "generated" by the four *Pauli spin matrices*:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{3.67}$$

These spin matrices are closely related to *quaternions* which play a great role in spatial rotations $O(3)$. This is another instance of the basic relation (3.66) (which is abstract and does not have an intuitive geometrical interpretation!).

The *unitary group in three complex* dimensions, $SU(3)$, is fundamental for the representation of hadrons in terms of *quarks*. Nobody can visualize such a complex space; it is a completely abstract mathematical tool. Nevertheless $SU(3)$, and later $SU(4)$ and $SU(5)$, have proved to be extremely efficient as order principles for the particle zoo.

$SU(3)$ has eight "generators", 3×3 matrices similar to (3.67), and was, therefore, by its discoverer Murray Gell–Mann (born 1929, Nobel Prize in 1969) called the "eight–fold way", half seriously after Buddha. (It should be mentioned that the co–discoverer of $SU(3)$ symmetry was Yuval Ne'eman.) An example is the *baryon octet* (8 particles) of Fig. 3.13. Here n and p denote neutron and proton; the other particles are strange baryons (hyperons), the superscript −, 0, + denote electrically negative, neutral, and positive.

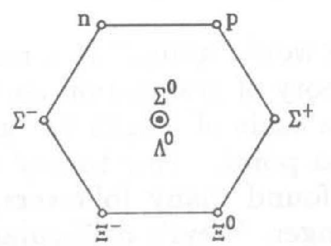


Figure 3.13: The baryon octet

The symmetry of Fig. 3.13 must not be taken literally in a simple geometric sense. This is also evident by the fact that there are 2 particles Λ° and Σ°, at the center! The correct expression for hadron symmetry is nothing simpler than the group $SU(3)$. If $\mathbf{U}$ is a transformation matrix pertaining to the group $SU(2)$ or $SU(3)$, the complex "vector" $\mathbf{x}$ on which it acts, and the resulting vector $\mathbf{y}$,

$$\mathbf{y} = \mathbf{U}\,\mathbf{x} \tag{3.68}$$

are called *spinors* in the case of $SU(2)$ and *3–spinors* in the case of $SU(3)$; they have 2 and 3 complex components, respectively. Enough!

Symmetry and the Greeks. The theory of symmetry played a great role in ancient Greece, starting at least with Pythagoras in the 6th century B.C. In the plane, symmetric figures are known to everyone: the equilateral triangle, the square, the regular pentagon, the regular hexagon shown in Fig. 3.13 etc. There are infinitely many regular polygons, one for every integer n.

In space this is quite different. There are no more than 5 regular polyhedra! They are called *Platonic solids* because, in his dialogue "*Timaeus*", Plato associates each regular polyhedron with an "atom" of the four "elements":

tetrahedron	...	fire
octahedron	...	air
cube	...	Earth
icosahedron	...	water

The most complicated regular solid, the dodecahedron, is, rather logically, considered the image of the universe as a whole (its face is the mysterious pentagon).

Depending on our relation to the history of our culture, we may smile at the naivité of the Greeks or admire their marvellous intuitive anticipation of the role of symmetry for elementary particles. I prefer the second attitude.

Gauge theories. The word "gauge" is a misnomer. In 1918, trying to establish a unified theory of gravitation and electromagnetism, Hermann Weyl introduced a scale of length (a "gauge" in the true sense) that varies from point to point. This theory was the first attempt at such a unification and found many followers, among them Einstein, Eddington, and Schrödinger. Weyl's particular approach was given up soon, however. Much more successful was Weyl's attempt in 1929 to find a theory of the electromagnetic field in interaction with charged particle fields. The *linear* change of scale was replaced by a *circular* change, a variable rotation, and the theory worked: the first gauge theory was found. It permitted an elegant derivation of Maxwell's equations for the *electromagnetic field*, which was thus recognized as a gauge theory for plane rotation, $O(2) \hat{=} U(1)$ in the terminology used above.

Weyl's principle remained forgotten for 20 years, until C.N. Yang and R.L. Mills developed a gauge theory for $SU(2)$ in 1954, which then turned out to be useful for the description of weak interactions. In 1960, Sheldon Glashow outlined a unified theory of electromagnetic and weak interactions, using the "product group" $SU(2) \times U(1)$ as the basis of a gauge theory. In 1967–1968, Steven Weinberg and Abdus Salam perfected the theory, which became known as the *standard theory of* *electro–weak interaction*. Glashow, Salam and Weinberg received for it the Nobel Prize in 1979: usually one gets the Nobel Prize for a theory only after it has been confirmed experimentally.

After Ne'eman and Gell–Mann had discovered the quark structure of hadrons also around 1960, described by means of $SU(3)$, it "only" remained to find a corresponding gauge theory for $SU(3)$ to describe the strong interaction. Then a gauge theory of $U(1) \times SU(2) \times SU(3)$ would be expected to furnish a "*grand unified theory*" (GUT), at least of three basic forces: *electromagnetic, weak, and strong interaction*. The theory thus obtained works so well in practice that is has been called *Standard Model*, but it contains some *ad hoc* assumptions which render it not yet perfect from a theoretical point of view.

By far the most difficult obstacle is the fact that, of the four basic forces, it is most difficult to incorporate *gravitation*. Ingenious theories have been elaborated: *supersymmetry, supergravity, string theory, superstring theory*. What seems to be missing, however, is a completely new idea from which a unified theory would follow automatically. Some scientists think that superstring theory might be such an idea, others are strongly against. (There is little doubt, however, that gauge theories, just like relativity and quantum theory, express a valid physical principle which will somehow have to be incorporated in any future "final" theory.) A fascinating account, providing an excellent introduction followed by interviews with leading proponents and opponents, is found in (Davies and Brown 1988).

This is the best general reference for this section. Books on particle physics become obsolete almost before they are published. So none of them is recommended here; the interested reader will easily find the newest books in bookstores or libraries. As general treatises involving also philosophical aspects (cf. sec. 6.6) we mention (Davies 1984), (Barrow 1991), (Weinberg 1993), (Kaku 1994), and (Gell–Mann 1994), each book having its own merits. On symmetry in general, (Weyl 1952) is still a classic; a very readable modern treatment is (Mayer–Kuckuk 1989).

3.7 Space and time; cosmology

Time is the moving image of eternity.
Plato

Space–time structure

As we have seen in sec. 3.4, the theory of relativity implies a unification of space and time. By the line element (3.45) of special relativity

$$ds^2 = -dx_0^2 + dx_1^2 + dx_2^2 + dx_3^2 \quad , \tag{3.69}$$

(p. 94) the spatial coordinates $x_1 = x$, $x_2 = y$, $x_3 = z$ are related to time t or the time coordinate

$$x_0 = c\,t \quad , \tag{3.70}$$

c being the velocity of light, but x_0 is distinguished essentially by the minus sign in (3.69). Thus ds^2 is not necessarily positive: we have

$$\begin{aligned} ds^2 &> 0 \qquad \text{on a space–like line ,} \\ ds^2 &= 0 \qquad \text{on a light–ray ,} \\ ds^2 &< 0 \qquad \text{on a time–like line .} \end{aligned} \tag{3.71}$$

This is illustrated by the space–time diagram of Fig. 3.14 and in more detail in Fig. 3.15. Time–like lines lie within the light cone; they represent possible *world–lines* of a person (from birth at A to death at B) or of a material object (mass point). Space–like lines are usual lines in space, for instance a plumb line, the edge of a ruler, or an electric wire.

These lines need not be straight, but for simplicity we shall use straight lines.

In Fig. 3.14, the "past" contains all *events* (points in space–time) which can influence the event P. The "future" consists of all events that can be influenced by P. The "present" contains all events which neither can influence P nor can be influenced by P since all effects propagate with velocity smaller than the light velocity c. By choosing x_0 as time coordinate we have implicitly put $c = 1$, which accounts for the angle of 45° in Figures 3.14 and 3.15. Thus the "present" looks so big.

In fact, however, c expressed in metric units is a great number, close to 300 000 km/sec, so that a realistic picture looks quite different

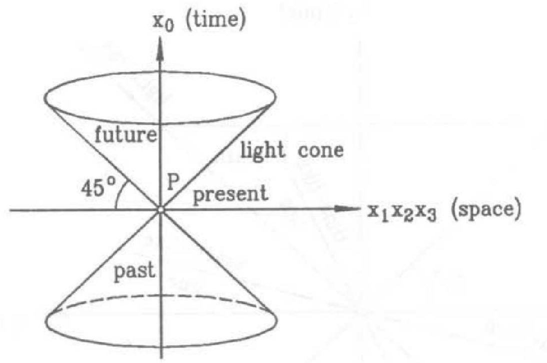


Figure 3.14: Space–time diagram

(Fig. 3.16). The realistic "present" is extremely narrow; it is an exact plane as $c \to \infty$.

The situation in the curved space–time is locally entirely the same, in the same way as a surface can locally be approximated by its tangent plane.

Creativity and the block universe. It has been claimed that space–time is a fixed four–dimensional manifold, in which an observer only "creeps" passively along his world–line from birth A to death B, being helplessly and irreversibly dragged along by time. This is the theory of the *block universe*: space–time exists ready–made, time is an illusion, the world simply *is*, it does not *become*. We are passive spectators, observing the flow of time without being able to influence it. Freedom of the will is an illusion. We need not worry about immortality: the world line from A to B in Fig. 3.15 simply *is*, birth, development, decay and death are all illusions. We are sitting in the cinema of life, watching the events which go on around us. We cannot change them, exactly as we cannot change the action going on in the television program we are just watching (we cannot even turn off the "TV set of time").

This picture is difficult to accept, but it cannot be refuted logically. If the line element of space–time were "positive definite",

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2 \quad , \tag{3.72}$$

with time x_4 *real* (in a mathematical sense) so that ds^2 can only be positive, the block universe would in fact be the only possible interpretation. There would be no light cone separating past and present; time x_4 would just be an ordinary space coordinate, of the same character as x_1 or x_2.

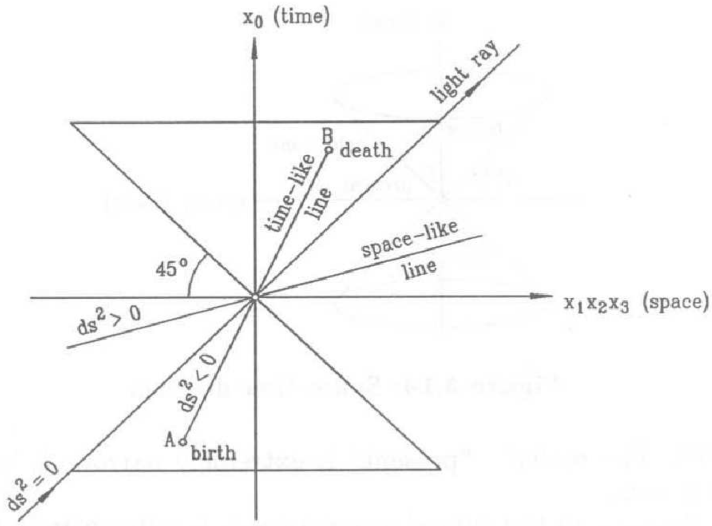


Figure 3.15: Various lines in space–time

It is the *minus sign* which makes all the difference, creating the light cone separating past, present and future. We can travel in space but not in time (see below). It is true that the Lorentz transformation (3.38) also transforms time, but the change is usually negligible in practice since $v \ll c$; cf. also the "realistic" third picture in Fig. 3.16.

Nevertheless, if we wish the flow of time to be real (in a philosophical sense), we should like time to have a unique direction, along which the creative process of life and mind can proceed. The "present" should consist of simultaneous events, which would require the two lines in the third, "realistic" model in Fig. 3.16 to coincide exactly, forming a single line separating past and present. In reality, however, the two lines in this figure form an extremely small angle, instead of coinciding as they do in classical mechanics: the present is "thin", but not "infinitely thin".

Whitehead and, more recently, Havemann (1964, p. 92) have pointed out a way out of this dilemma. Past and future must be redefined in agreement with relativity: past is of what we already can have knowledge, future is what we can still influence by our actions. Thus *creative action* is possible. Time is not an illusion, history really goes on, biological evolution exists, animals and men are born, live, and die. We are actors and not merely passive spectators in the universe. Let us call this model *open universe*.

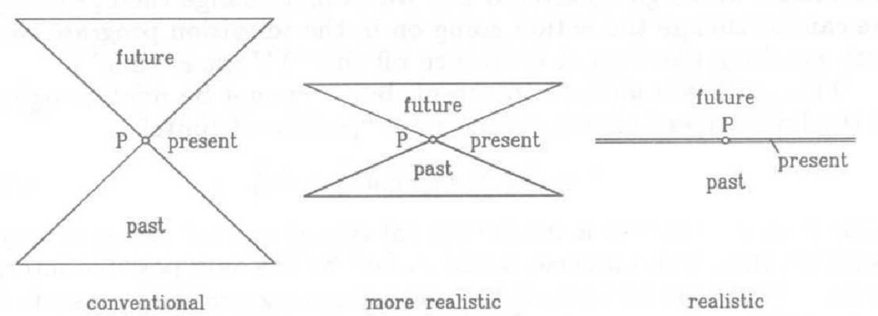


Figure 3.16: Conventional and "realistic" light cones

Block universe and open universe contradict themselves only if taken as absolute models. *Regarding the past, our universe is a block universe*: the past is past and can no longer be changed. *Regarding the future, the universe is open*. Now, in 1994, we are open towards the year 2000, and can take active and creative steps to ensure that mankind still exists in 2000; in 2001, the year 2000 is in the past and can no longer be influenced. For 2001, the universe up to and including the year 2000, is a block universe.

Again we have a Hegelian synthesis: the creative open universe is the thesis, the block universe is the antithesis, and the distinction: block universe for the past, open universe for the future, may be the synthesis.

We also may say that the block universe represents the "static" geometric view, whereas the open universe expresses a creative "dynamic" view.

Cosmic time and the evolution of the universe

Locally, general relativity behaves exactly like special relativity, with the light–cone separating past, present, and future. The global structure is a block universe, but an open universe with "real" creative time is consistent with general relativity as well.

It is even more so because the basic model universes of general relativity admit a *uniquely defined distinguished direction of time*. They have a general form (*Robertson–Walker metric*; "metric" is the same as "line element"):

$$ds^2 = -c^2dt^2 + \left(\frac{R(t)}{R_0}\right)^2 (dx^2 + dy^2 + dz^2) \quad , \qquad (3.73)$$

for a flat three–dimensional space (curvature 0); for a space of constant curvature (positive or negative) the second term on the right–hand side is slightly more complicated. (Curvature of 3–space is essential for any rigorous treatment; our disregard is motivated only by reasons of didactic simplicity.)

Here R_0 is an irrelevant scale factor. In the special case $R(t) = R_0 =$ const., eq. (3.73) reduces to the flat space–time of special relativity. Generally, $R(t)$ may be intuitively regarded as the *radius of the universe*, although this is another barely acceptable oversimplification; cf. (Misner et al 1973, p. 721).

Without going into details, we can distinguish three possibilities, shown in Figures 3.17 to 3.19. The interpretation is obvious. The static model of Fig. 3.17 corresponds to a universe that does not change with time. It is excluded by the empirical fact that the universe is expanding. Remote galaxies are moving away from us (or we are moving away from them) with speeds which are the greater the farther away the galaxy is. The mysterious *quasars* which are extremely far away, are moving away with a speed approaching that of light.

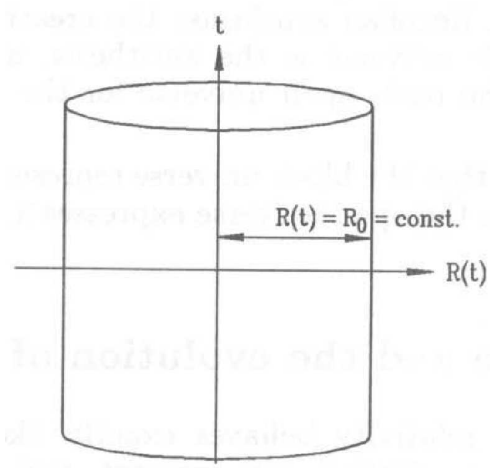


Figure 3.17: A static model

Fig. 3.18 is regarded at present to be the most realistic model. The universe started, some 13 billion years ago, as a point singularity A. This would correspond to the creation of the world, but there is a growing tendency in science towards sober and even humorous understatement (think of the name "quark" for the most important

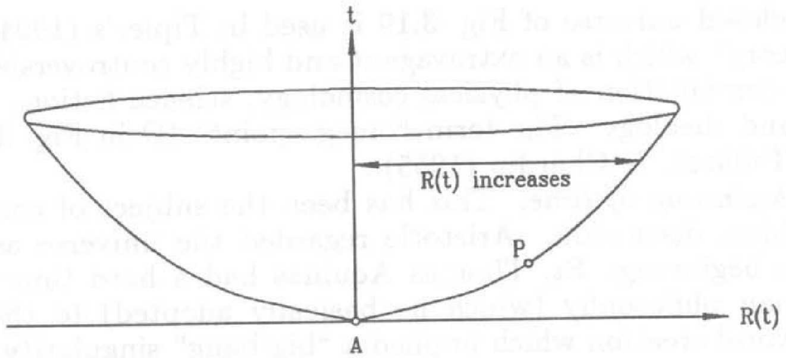


Figure 3.18: An expanding universe

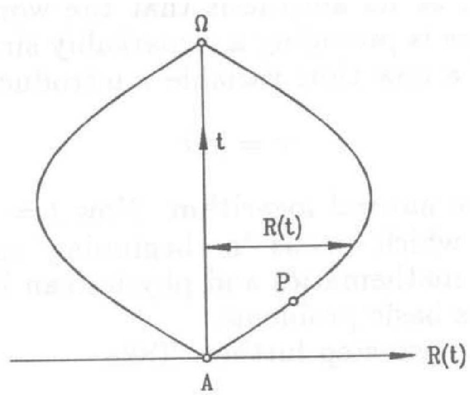


Figure 3.19: A closed universe

sub–elementary particle); therefore physicists prefer the term *big bang*. After that, $R(t)$ increases: *the universe expands*.

Fig. 3.19 shows a model which starts with a big bang singularity at A. The radius $R(t)$ first increases: the universe expands, but then it contracts again to another pointlike singularity Ω.

At present, we appear to be at a point P, at a phase of expansion. It had not been decided whether our actual universe corresponds to Fig. 3.18 or 3.19. This question can be settled, at least in principle, by determining the mean density of the universe. This is very difficult and controversial, however, mainly because of the existence of invisible "dark matter" in the universe. The majority of cosmologists seem to favor the *expanding universe* of Fig. 3.18.

The closed universe of Fig. 3.19 is used in Tipler's (1994) "omega point theory" which is an extravagant and highly controversial but fascinating combination of physical cosmology, science fiction, computer theory, and theology. The term "omega point" (Ω in Fig. 3.19) goes back to Teilhard de Chardin (1955).

The beginning of time. This has been the subject of considerable philosophical discussion. Aristotle regarded the universe as eternal: it has no beginning. St. Thomas Aquinas had a hard time adapting Aristotelian philosophy (which he basically adopted) to the biblical view of world creation which implied a "big bang" singularity. See also St. Augustine's view in sec. 5.4.

Kant's first antinomy states as a thesis that the world has a beginning in time and as its antithesis that the world has been lasting forever. Mathematics is providing a remarkably simple solution of this antinomy. Think of a new time variable τ introduced by

$$\tau = \ln t \tag{3.74}$$

where ln denotes the natural logarithm. Now $t = 0$ (big bang) corresponds to $\tau = -\infty$, which means "no beginning" in terms of τ!

This shows that mathematics and physics can indeed help philosophy solve some of its basic problems.

We can even take one step further. Take

$$r = \sqrt{x^2 + y^2 + z^2} \tag{3.75}$$

as a radial coordinate. Then we may put in (3.73)

$$R(t) = c\,t \quad , \tag{3.76}$$

which corresponds to linear expansion according to a particularly simple model proposed by the British cosmologist E.A. Milne. A particular galaxy moves away with constant velocity v:

$$r = v\,t \quad . \tag{3.77}$$

Now introduce a new radial variable ρ by

$$\rho = \operatorname{arth}\frac{r}{ct} \tag{3.78}$$

where "arth" denotes "area tangens hyperbolicus", a well–known elementary mathematical function.

Now comes the surprise: for our galaxy, eqs. (3.77) and (3.78) give

$$\rho = \mathrm{arth}\frac{v}{c} = \mathrm{const.} \tag{3.79}$$

and for the "limit sphere"

$$R = r_{max} = c\,t \tag{3.80}$$

which is the maximum r possible at time t since $c = v_{max}$, light velocity as the maximum possible velocity. Now (3.79) gives

$$\rho_{max} = \mathrm{arth}\frac{c}{c} = \mathrm{arth}1 = \infty \quad . \tag{3.81}$$

That means, all ρ's of galaxies is constant, and space is infinite.

Thus simple mathematical equations transform an expanding universe of the type of Fig. 3.18 to a static model of type Fig. 3.17 (with $R = \infty$).

But the expansion of the universe has been confirmed by observation! By which observation? By means of the *Doppler shift* of spectral lines. If a police car, with its siren howling, approaches me, the sound appears higher than when it has passed me and happily disappears in the opposite direction. Thus for an approaching object, the frequency appears higher than normal, and for a receding object, the frequency appears lower. For light, red light has a lower frequency than blue light, so the light of an approaching star would appear shifted towards blue. What we observe, however, is a *red shift* which indicates that the stellar object (star, galaxy, quasar) moves *away* from us, and the greater the redshift, the greater the speed. This is the observation method by which the expansion of the universe is measured.

But a frequency change could also be explained by assuming that in the past, light has had a lower frequency than today. As we observe more and more distant stars and galaxies, we also go back in time. A star at a distance of one light year is observed as it was a year ago, because light took one year to travel from the star to the observer. If we observe a galaxy which is a million light years away, we see it, not as it is now, but as it was a million years ago! Hence, if the observed galaxy has a strong redshift, it means that it is far away in time (since it is far away in space) and we see it as it was long ago when the clocks (oscillating atoms or molecules are clocks!) were much slower than they are now. Mathematically, this means that time is measured in terms of τ rather than of t, and logically there is no difference between explaining redshift in terms of an expanding dynamic universe and explaining it in terms of increasing "speed of natural phenomena". (In fact, it seems that the speed of modern life appears much greater than that of life in the "good old days" — please don't take this argument seriously!)

Although most astronomers and physicists prefer describing the universe in terms of r and t to the description in terms of ρ and τ, we see that "things ain't that simple". So a good knowledge of mathematics may sometimes help regard philosophical controversies (beginning of time, finite or infinite universe) with less emotion.

It is similar with the classical Greek *paradox of Achilles and the tortoise*: though he is much faster than the tortoise, Achilles can never reach it. Let us assume that Achilles is a hundred times faster than the tortoise, and initially they are separated by 1000 meters. After Achilles has covered this distance of 1000 m, the tortoise has moved 10 m, after Achilles has covered these 10 m, the tortoise has moved 0.1 m and so on. If both Achilles and the tortoise are pointlike, before Achilles has reached the animal, the latter has always moved a little farther: Unlucky Achilles will never reach the smiling beast. Modern mathematics reduces the problem to the summation of the infinite series

$$\begin{aligned} &1000\,\mathrm{m} + 10\,\mathrm{m} + 0.1\,\mathrm{m} + \cdots = \\ &= 1000\,\mathrm{m}\left(1 + \frac{1}{100} + \frac{1}{100^2} + \frac{1}{100^3} + \cdots\right) = \\ &= 1000\,\mathrm{m}(1 + q + q^2 + q^3 + \cdots) = \frac{1000\,\mathrm{m}}{1-q} \\ &= \frac{1000\,\mathrm{m}}{1-\frac{1}{100}} = \frac{100000}{99}\,\text{meters} \quad . \end{aligned} \tag{3.82}$$

This is the distance at which Achilles overtakes the tortoise, to convert it into a tasty turtle soup.

This example, which looks simple to us now, has, for some 2000 years, played a serious role in philosophic discussions about time!

Gödel's universes and time travel

The logician Kurt Gödel, of Austrian origin but then, together with Einstein, at the prestigious Institute of Advanced Studies at Princeton, wrote perhaps 5 or 10 papers throughout his life, but each creating a major sensation. In the Einstein Festschrift (Schilpp 1949), Gödel published on pp. 555–562 a paper "*A remark about the relationship between* *relativity theory and idealistic philosophy*". There he described a mathematically perfectly valid cosmological solution of Einstein's equations of general relativity which *contained closed time–like world lines*.

Usually, world lines are not closed, cf. the line AB in Fig. 3.15 on p. 122. Closed time–like world lines permit you to travel into your own past (Fig. 3.20). This may have paradoxical consequences. You travel from P via Q to point G where you kill your grandmother — not intentionally as sometimes claimed in the literature, but by accident when, as a small girl, she was playing in the street and you, driving too fast, saw her only when it was too late. Anyway, no grandmother, no mother, no ... you!

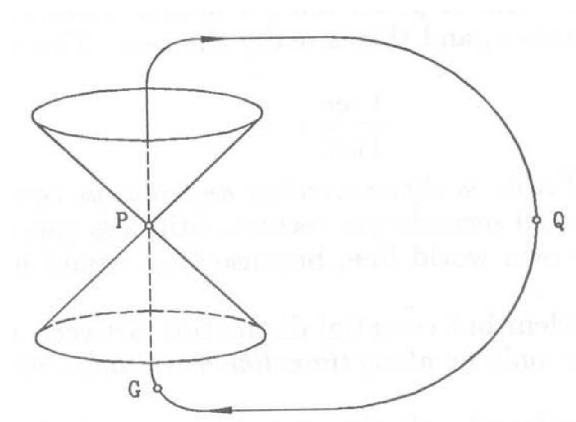


Figure 3.20: A closed time–like world line

Paradoxical conclusions of this type may occur with closed time–like lines. Hence, Gödel concluded that space–time must be "unreal": the passing of time is a mere illusion, and the universe is a *block universe* (see above). In his reply to Gödel (Schilpp 1949, p. 688), Einstein was worried and cautiously remarked: "It will be interesting to weigh whether these [solutions found by Gödel] are not to be excluded on physical grounds."

In his excellent book, Reichenbach (1957, pp. 141–142) gave an even more interesting example: you meet a man who claims to be your younger self. This idea is used "iteratively" in an irresistibly funny way in the Seventh Voyage of Ijon Tichy in the science fiction book "*The Star Diaries*" by Stanislaw Lem. The space ship, going through strongly irregular gravity fields of Gödel type, keeps filling and crowding with identical Ijon Tichy's who are multiplying by going through time loops of one day's duration: the Ijon Tichy of Monday quarrels with the Ijon Tichy of Tuesday and that of Wednesday and so forth.

The first major tale of time travel is the classical novel "The Time Machine" by H.G. Wells published in 1895, where he anticipated several ideas of the theory of relativity, especially the similarity of time and space.

He made, however, a philosophically very important error. Taking the identity of space and time too literally, his hero travels into the future much more rapidly than the "ordinary rate of time". Now, what is the rate of time flow? What does it mean? What is the "speed of time"? Ordinary "spatial" velocity is measured in meters per second, m/sec, say (automobilists prefer km per hour). Thus, the "velocity" of time would be sec/sec, and this is really the case. The velocity of time, by definition, is

$$\frac{1\,\text{sec}}{1\,\text{sec}} = 1 \quad , \tag{3.83}$$

the "velocity" of time is dimensionless and always equal to unity ! It simply cannot be 20 seconds per second. Still less can we travel backwards along our own world line, because that would involve negative "velocities".

This is an evident but essential distinction between space and time. "Time travel" can only be *along time–like world lines with "speed" equal to unity*.

All about time travel – physics, metaphysics and science fiction – can be found in the comprehensive work (Nahin 1993) which also contains an incredible number of references. Very nice is also (Rucker 1984).

The arrow of time

Why does time flow in one direction only? Why are we irresistibly getting older, instead of being able to return to our youth which many people (excluding the present author) wish?

The equations of mechanics, classical relativistic and even quantum mechanics, are *reversible* in time, that is, they keep their validity if t is replaced by $-t$.

There are, however, irreversible processes as we have seen in statistical mechanics and in the measuring process of quantum mechanics, see secs. 3.2 and 3.5. They somehow seem to involve statistics or, what is the same, probability.

Consider, for example, a box consisting of two chambers, chamber A filled with a gas, and B empty. The shutter S between them is closed.

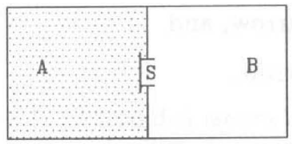


Figure 3.21: The left chamber is filled by a gas, the right chamber is initially empty

Now open the shutter: gas will flow from A to B until the gas is equally distributed in A and B. If we assume that the gas molecules move according to the reversible laws of classical mechanics, the opposite process should also occur, at least from time to time: the whole amount of gas now in B should flow back into A, and B should be empty, at least for a moment.

This should happen, but it never happens. The explanation is statistical: the back flow *does* happen, but so seldom that there is no chance whatsoever to watch such an event. The probability of such an event is so small that it is practically zero.

To see what a time–reversal means for our practical life, run a motion–picture backwards. Autumn leaves fall upwards, glass splinters put themselves together to form a new bottle, two cars badly damaged by a collision separate and become nice undamaged automobiles. The biographical film of a famous man starts by showing him dead in his grave. He resurrects from the grave and becomes progressively younger. He goes to university, then to secondary school, to elementary school, then to kindergarten. He becomes a baby and loses most of his hair. His teeth become smaller and smaller and finally vanish. The end of the film is an ugly little thing in a cradle, admired by his mother and, somewhat more hesitantly, by his father.

What is the explanation of the arrow of time? I think, ultimately we just have to accept it, just as we accept ourselves and our surroundings. Nevertheless, explanations will help get a deeper understanding of this mysterious phenomenon, which only looks natural because we have got used to it.

So to speak, we have five main "arrows of time":

– the thermodynamical arrow,

– the biological arrow,

– the historical arrow,

– the psychological arrow, and

– the cosmological arrow.

The *thermodynamical arrow* is based on the remarkable irreversibility of statistical mechanics as described above (Fig. 3.21). It is characterized by the steady increase of *entropy* (sec. 4.3).

The *biological arrow*, in a sense, has opposite character although it points along the same direction. It is the arrow of biological life, from birth to death. It is the arrow of evolution, from amoeba to man. It means increase of biological information contained in the genes.

The *historical arrow* is closely related. It is characterized by the increase of *our* information. The unknown future becomes present and then history, to be studied by historians, archeologists, and paleontologists. The increasing knowledge is stored in libraries of rapidly increasing size. Universities are growing and becoming diversified at a breath–taking pace.

The *cosmological arrow* corresponds to the expansion of the universe. There is also a cosmic evolution, starting from a probably rather undifferentiated "fireball" of unimaginably concentrated energy after the "big bang" and leading to galaxies, stars and planets. It is accompanied by a chemical evolution from hydrogen to all the heavier chemical elements; cf. (Weinberg 1977).

The *psychological arrow of time*, our subjective feeling of the passage of time, is related to the biological arrow, to the historical arrow (the accumulation of information in our brains), and possibly even to the thermodynamical arrow: a computer processes information at the expense of energy (high–quality electric energy is degraded into low–quality thermal energy which must be removed by a cooling system), and the brain should work in a somewhat similar way (Hawking 1988, p. 147).

Wheeler's example. In his Nobel–Prize lecture, R.P. Feynman talked about a telephone conversation between the two great physicists (we are following (Gardner 1982, p. 268)):

> "Feynman", said Wheeler, "I know why all electrons have the same charge and the same mass."
> "Why?" asked Feynman.
> "Because", said Wheeler, "they are all the *same* electron."

What does this mean? Consider Fig. 3.22, following Whitrow (1980, p. 332). The picture on the left admits the following standard interpretation. At A, a photon spontaneously "splits" into an electron and a

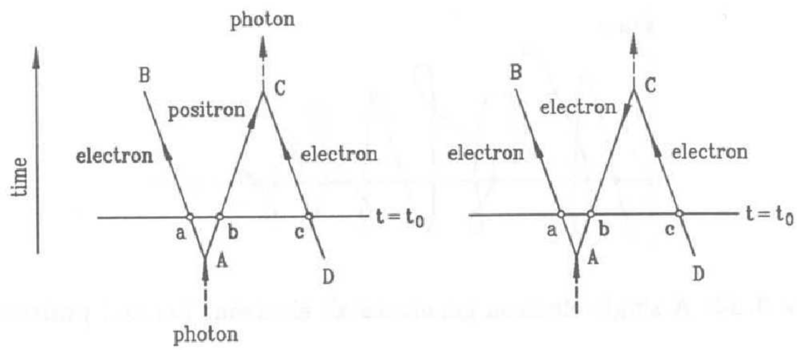


Figure 3.22: A positron is an electron running backward in time

positron (energy and electric charge are conserved!). At C, the positron meets another electron; the two particles annihilate each other, emitting a photon. According to Wheeler's interpretation shown in the picture on the right, an electron starts at D, collides with a photon at C, suffers a "recoil" which sends it back in time to A, where it collides with another photon which makes it change direction in time again, sending it to B.

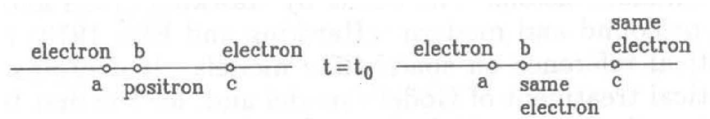


Figure 3.23: Two electrons and a positron or one and the same electron at three different places?

At a certain instant t_0, we thus have the picture shown in Fig. 3.23. Thus, if *we regard a positron as an electron running backward in time*, we have, at the same time $t = t_0$, the same electron in three different places a, b, and c (Fig. 3.23, right)!

This can be iterated. Fig. 3.24 shows the world–line of a single electron. It generates all electrons and positrons in the world.

Wheeler's idea is extremely interesting, but it cannot be entirely true: it would imply that there are as many positrons in the world as there are electrons. This is manifestly not the case.

The philosopher A.N. Whitehead once remarked: "It is more important that a proposition be interesting than that it be true."

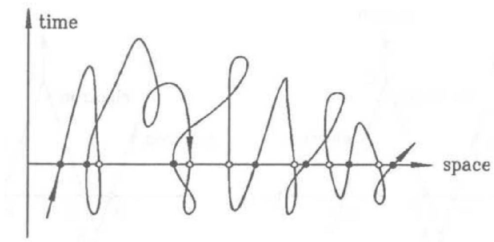


Figure 3.24: A single electron generates all electrons (•) and positrons (∘)

Wheeler's idea is so beautiful that it almost has to express a true partial aspect of nature. Many physicists now think that the arrow of time is a macroscopic "global" feature of the world. On a microscopic, quantum–mechanical, "local" scale, time may well be reversible.

Additional reading. The basic reference on time is (Whitrow 1980), unsurpassed in comprehensiveness, depth and readability. (Gardner 1982), and (Rucker 1984) are "popular" books of high level. (Čapek 1976) and the smaller book (Smart 1964) are anthologies of articles of authors from Greek antiquity to the present time. (Reichenbach 1957) is a very readable classic. The books by Hawking (1988 and 1993) are readable, profound and modern. (Hawking and Ellis 1973) is the basic mathematical reference on space–time models. It contains a detailed mathematical treatment of Gödel's model and, for the first time, shows a very instructive geometrical picture, cf. their Fig. 31 on p. 169. Once more we mention (Treder 1974).

3.8 Inverse problems

Most mathematical problems in science,
technology and medicine are inverse problems.

Gottfried Anger

Introduction

In a poem, addressed to a physicist, Johann Wolfgang von Goethe ascribes to him the opinion (which he himself rejects):

Ins Innre der Natur
Dringt kein erschaffner Geist
Glückselig, wem sie nur
die äussre Schale weist!

("No created spirit penetrates into the interior of nature. We are already lucky if nature shows us only its external shell.")

This is exactly the situation which we are facing in *geophysics* if we want to investigate the *interior of the Earth*. Our measurements are restricted to the Earth's surface, and we want to use them for finding out the Earth's interior structure. Such measurements are, above all, observations of seismic waves originating from large earthquakes, but also, e.g., observations of the Earth's external gravitational field.

Physicians face very much the same problem if they want to investigate the *interior of the human body* , for instance in order to discover a tumor. Of course, they can directly penetrate into the interior by surgery, and they will do so if they know where the tumor is. But first they will use indirect methods such as X–rays (which closely correspond to the seismic waves of the geophysicist) or nuclear magnetic resonance (NMR) tomography.

Example 1. Consider the Earth's gravitational field. Obviously, the gravitational force is produced and determined by the distribution of the masses of density ρ inside the Earth. Symbolically we may write

$$g = A\,f \tag{3.84}$$

where g denotes the gravitational force at the Earth's surface, $f = \rho$ represents the distribution of density ρ inside the Earth, and A is the "*Newton operator*" (a generalization of Newton's law (3.11) on p. 74) computing g if f is given. ("Operator" is here again used in the sense of performing a mathematical operation.) An explicit form of the Newton operator will be found in sec. 6.6, eq. (6.2) on p. 268.

Let us now consider g as given by gravity measurements; the problem is to determine the density distribution f inside the Earth.

Formally, the solution is simple:

$$f = A^{-1}g \tag{3.85}$$

where A^{-1} denotes the *inverse* to the Newton operator. It is a task of mathematicians to find it; this task is not impossible, but still surprisingly difficult if we consider the simplicity of the *direct problem* (3.84). Eq. (3.85) then expresses a solution of the *inverse problem*.

Symbolically, we may express the problem by the diagram of Fig. 3.25. The operator A is a "*black box*" which converts the input (density f) into the output (gravitation g). This is the *direct problem*.

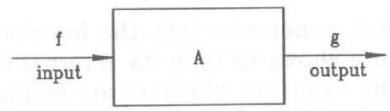


Figure 3.25: An input–output diagram

The diagram behaves well if we follow it from left to right, in the sense of the arrows, that is for the direct problem. For the *inverse problem* (3.85), we have to go through the diagram in the opposite direction, from right to left, against the direction of the arrows. This causes difficulties.

The general structure expressed by Fig. 3.25 and equations (3.84) and (3.85) also applies to many other cases. The operator A may be *linear* (sec. 3.5) or *nonlinear* (sec. 3.2). The Newton operator is linear.

Example 2. Projection. Consider a photographer A taking the picture g of a pretty girl f. This example is immediately clear, and it is directly relevant to the problem of *imaging*. Thus f denotes the *object* (also the interior, e.g., in X–ray pictures), and A represents the *projection* producing the *image* g. The operator A may denote amateur photography, but also an ordinary X–ray equipment, X–ray tomography or NMR (nuclear magnetic resonance) tomography in medicine, as well as seismic tomography in geophysics. In any case, g will be the output data.

This mathematical model thus is very general; the operator A is now called a *projection operator.*

Example 3. Human perception. This is a non–mathematical analogue of Example 2. Here, the "equations" (3.84) and (3.85) must be taken in a very general sense, as a symbolic expression of the diagram of Fig. 3.25.

The symbol "f" denotes an *object*, or *reality*, or *nature*. The "projection operator" A symbolizes *sense perception*, by means of the eye or the ear, and the "output" g are the *sense data*, cf. sec. 1.3. To emphasize the active role of perception ("searchlight" rather than "bucket", cf. sec. 1.4), A may also be called a *search operator*.

The "direct problem" thus is *sense perception* of nature. The corresponding "inverse problem" is the determination of f (object, reality, nature) from sense data g. This is the basic problem of the *theory of knowledge*, or *epistemology*. There is no need to point out its difficulty, which has kept philosophers busy from Socrates and Plato to Karl Popper.

Classification. Our mathematical (or pseudo–mathematical) symbolism permits a useful classification:

given	to be determined	classification
A, f	g	direct problem
A, g	f	inverse problem of first kind
f, g	A	inverse problem of second kind

Before we discuss these problems in more detail, we must define "well–posed" and "ill–posed" problems.

Ill –posed problems

A problem is called *properly posed*, or *well–posed*, if the solution satisfies the following three requirements:

(1) existence,
(2) uniqueness,
(3) stability.

This means that a solution must exist for arbitrary (within a certain range) data, that there must be only one solution, and that this solution must depend continuously on the data. If one or more of these requirements are violated, then we have an *improperly posed,* or *ill–posed problem.* For a long time it was thought that only properly posed problems are physically meaningful. In fact, deterministic processes, as considered in classical mechanics, depend uniquely and continuously on the initial data — this is the essence of causality — and thus correspond to properly posed problems.

Only relatively recently it was recognized that there are important problems that are not properly posed. There is now an extensive literature on improperly posed problems; cf. (Anger et al. 1993).

In fact, *most inverse problems are ill–posed*, an extreme example being epistemology (or philosophy in general). But even some direct problems are ill–posed, as we know from weather prediction and, more generally, from chaos theory (sec. 3.2).

Physically, well–posed problems are *stable*, and ill–posed problems are *unstable* . Laplace's demon, embodying classical determinism and causality, operates in a stable fashion; St. Peter (if regarded as the saint responsible for weather) has a rather unstable character.

Inverse problems of first kind

As we have seen above, such problems consist in solving the equation $g = Af$ for f. A formal solution has been given by eq. (3.85).

In this inverse problem, g represents the measured *data*, and f represents parameters that describe the *object*.

The simplest and most important case thus is provided by *Example 2: Projection.* In this case, the data g are some "image" of "nature" f, the projection being defined by the operator A. This case ranges from amateur photography to tomography in medicine and geophysics.

Thus, generally, the *projection operator* A works in the direction from nature to observation, it projects from "object space" into "observation space". It must be inverted to give A^{-1}, working from observation to nature and determining some feature of nature or reality f from observations g. Many alternative formulations may be given, as we have seen before. Since the apparatus A (X–ray, tomography) is searching to get information g from nature f, A has also been called a *search operator* to express its active role.

Let us try to understand this better by means of the concrete example of *tomography.* Here the objective is to determine the *inner structure* of the object f. In medicine we have X–ray and NMR tomography; in geophysics we have seismic tomography. The mathematical structure is similar in all three cases. The data g comprise a set of X–ray images or NMR data or observations of seismic waves.

Now we shall return to our *Example 3: Human perception.* As we have already mentioned above, sense data g provide information about reality f. The projection operator A here is not an X–ray or tomographic apparatus but the human organs of sense perception such as the eye and the ear, plus the additional "hardware" as provided by our nervous system which has been developed (at least to a large extent) by biological evolution. Continuing this analogy, we may say that experience and learning have improved or even created the corresponding "software" which may be considered some modern equivalent of Kant's *a priori.* This "evolutionary theory of knowledge" has been described in sec. 1.4.

In this context, calling A a *projection operator* corresponds to Russell's theory of sense data acquired passively and then analyzed, whereas the name *search operator* for A corresponds to Popper's theory of human perception as an active "searchlight". (This is really a gross oversimplification justifiable, if at all, then only by its didactic purpose.)

Of course, human perception is not directly a mathematical problem: the operators A and A^{-1} are realized by biology or physiology, with a built–in mathematical structure and the corresponding "neuro-computational" systems.

Inverse problems of second kind

Above we have defined inverse problems of the second kind as the determination of the operator A, considering both f and g as given data. The basic mathematical structure is, of course, always $g = Af$.

Here it is useful to consider the operator A as a *law* which transforms certain given or measurable parameters f into other measurable parameters g. Hence f and g can be prescribed or measured; the law A is then to be determined.

For the law A we may have a certain number of alternatives; in simple cases this reduces to a *statistical testing of hypotheses*. Or we essentially know the law, and only a few parameters are to be determined. A particularly interesting case is Newton's derivation of his law of gravitation A from the Keplerian laws of planetary motion, which thus served as data for determining the "grand" law A.

Generally matters are not quite that simple. Laws are usually found by "guessing" or, using a more respectable expression, by physical intuition. Let A be such a hypothetical law; it must now be tested by comparing it with experimental data f and g: this is a *verification* of A (there must be $g = Af$). It may "survive" all known tests; is this sufficient to assert that it is true? This is the problem of *induction*.

Well, we can never be sure that another verification may not show $g = Af$. Take the example of the "law" that the sun rises in the morning and sets in the evening. This has been thoroughly confirmed by mankind; there was never a single exception. Is it thus necessarily true also for tomorrow? True with overwhelming probability, yes, but not with necessity: The Sun may have exploded during the night and may also have destroyed the Earth in this process.

Some contemporary philosophers of science, such as Sir Karl Popper, have replaced verification by *falsification*. Since verification can never be certain, it would be better to try to disprove the theory by all possible means; if this is not successful (if one always gets $g = Af$), then we may, for the time being, accept the theory.

Matters are again not so simple, however. Because of measuring errors, $g = Af$ will never hold with absolute accuracy. A "neighboring" theory may hold as well within the limits of the accuracy of the experiment.

For instance, the theory of special relativity reduces to classical mechanics for velocities v that are small as compared to the light velocity c (in other terms, for $c \to \infty$), and quantum theory reduces to classical mechanics for $h \to 0$ where h is Planck's constant. Thus, in describing ordinary ("macroscopic") physical experiments, classical mechanics is sufficient, but if we absolutely wish, we may describe the experiments in a more complicated way also by special relativity or by quantum mechanics: the results will be practically the same.

Thus experiments must be carefully and ingeniously devised in order to distinguish between two theories, the so–called *crucial experiments*. An example is the well–known Michelson–Morley experiment for special relativity. It really permits to "falsify" classical mechanics in an extreme case, cf. secs. 3.4 and 3.9.

Also the gravitational field of the Earth and of planets may be correctly described by classical mechanics, if necessary with very small "relativistic corrections", although the description by Einstein's theory of general relativity is theoretically superior. This is a typical case of a phenomenon pointed out by the well–known mathematician Henri Poincaré: several different laws may fit equally well (if necessary with small corrections). Thus the choice is "conventional" and may be done by "esthetic" criteria such as "simplicity" or "mathematical elegance". This is called *conventionalism*.

The importance of measuring errors or "noise" in such considerations is evident.

Thus the theory of inverse problems may provide a preliminary first introduction to the theory of induction, verification and falsification which will be considered in more detail in the following section. Conventionalism will play a certain role in sec. 6.5.

Our standard *Example 3*: *Human perception*, may also be considered from the point of view of inverse problems of second kind. Here the operator A is the apparatus of human perception, called by Konrad Lorenz (1973) the "mirror" ("Spiegel") which mirrors our environment for us. This "mirror" is human perception including neural "hardware" and "software", cf. secs. 1.2 and 1.3. Thus the investigation of this "mirror", looking with Konrad Lorenz at the "back side of the mirror" (die Rückseite des Spiegels), may be regarded as an inverse problem of second kind.

Thus "unreflected" sense perception is related to an inverse problem of first kind, as we have seen above. On the other hand, physiological investigation of and philosophical reflection about the mechanism of perception leads to an inverse problem of second kind.

System identification. Sometimes a mathematical or technological system is really a "black box" A in the sense of Fig. 3.25, and we wish to determine its internal structure "from the outside", without opening the "box". This is relatively simple if the general structure of the system is known and only a few parameters are to be determined. They characterize the materials by which the system is physically realized and are therefore called *material parameters*. For this purpose, a given input f is fed into the system, and the corresponding output g is measured. This is repeated for several f and g. In this way, equations may be obtained from which the parameters of the system A can be determined. A correct determination of material and other system parameters is of basic importance for sensitive systems which have to work reliably, for instance, aircraft engines. The determination of these system parameters is called *system identification*. The construction of a model of a complicated system is called *reverse engineering*.

Neural networks and learning. Such networks are attempts to provide a simple mathematical model for the activity of neurons in the human brain (sec. 1.1). A linear neural network represents an output signal g_i as a linear combination of the input signals $f_1, f_2, \ldots, f_n$:

$$g_i = a_{i1} f_1 + a_{i2} f_2 + \ldots + a_{in} f_n \quad . \tag{3.86}$$

The coefficients a_{ij} may be considered some kind of *weights* for the influence of the input f_j on the output g_i. If we have n such equations ($i = 1, 2, \ldots, n$), then we may write explicitly:

$$\begin{aligned} g_1 &= a_{11} f_1 + a_{12} f_2 + \ldots + a_{1n} f_n \quad , \\ g_2 &= a_{21} f_1 + a_{22} f_2 + \ldots + a_{2n} f_n \quad , \\ &\vdots \\ g_n &= a_{n1} f_1 + a_{n2} f_2 + \ldots + a_{nn} f_n \quad , \end{aligned} \tag{3.87}$$

or in vector–matrix notation using boldface symbols:

$$\mathbf{g} = \mathbf{A}\,\mathbf{f} \quad . \tag{3.88}$$

The problem is to determine the weight matrix $\mathbf{A}$ from given inputs $\mathbf{f}$ and their corresponding prescribed outputs $\mathbf{g}$. Since $\mathbf{A}$ contains n^2 unknowns a_{ij}, we need n given input vectors $\mathbf{f}^k$ and corresponding given output vectors $\mathbf{g}^k$, $k = 1, 2, \ldots, n$.

By determining $\mathbf{A}$ and realizing it physically or computationally, the system provides the required reaction $\mathbf{g}$ to the input $\mathbf{f}$. In other terms, the system has "*learned*" to produce the desired reaction $\mathbf{g}$.

Thus *learning* is formally reduced to system identification or reverse engineering, cf. (Spies 1993, pp. 311–313). *Learning is an inverse problem*!

More details on neural nets, especially the details of the determination of $\mathbf{A}$ and the introduction of non–linearities, may be found in the literature on neural networks; cf. (Kohonen 1988) and (Spies 1993).

Applications to psychology. Let the system A be a person applying for a job. He or she has to undergo a psychological test: questions f are posed and answers g are obtained. Thus the knowledge of the person A, his (her) reactions to stress, emotional stability, etc., generally his (her) qualification for the job are tested. It is of basic importance that the test (the input f) is so designed as to provide a balanced and objective information on the person A. This, of course, is a problem of psychological capability and experience, combined with common sense, of the person designing or directing the test, asking the right questions f, etc.

You might also test the character of a friend A in this way, observing his (her) reaction g to your behavior f. Do this tactfully, however, otherwise you may lose a friend.

Examinations have the same character: the professor puts the questions f to the student A and listens to his answers g. The only "system parameter" to be obtained in this way is the *grade* characterizing the student's performance in the examination.

These psychological applications share, with mathematical inverse problems, the logical structure, but the use of mathematics in this field is naturally rather limited.

Overdetermined and underdetermined problems

According to *classical causality*, the causes are necessary and sufficient to fully and uniquely determine the effects. Thus we may have n causes and n effects: the system is fully *determined*.

In *quantum mechanics* (sec. 3.5), the state after measurement is not fully determined by the state before measurement: the system is *underdetermined*.

In Whitehead's (1933, Chapters XII and XIV) *theory of actual occasions*, the past does not blindly and automatically determine the future as it does in classical causality. Especially in a mental event going on in one's consciousness, there are many incompatible and conflicting "causes" fighting for realization. Think of what is going on in your mind when you have to take a difficult decision. This has also a counterpart in the neural structure of the brain: many nerve impulses arrive, some excitatory and some inhibitory (sec. 1.1), and the total effect will imply some internal adjustment. This seems to be an *overdetermined problem*. Sometimes there may also not be enough data to automatically assure a unique solution; thus leaving room for the additional element of *creativity*.

It thus seems that we meet in nature underdetermined and overdetermined problems as well as the well–determined problems of classical mechanics.

A simple mathematical model for overdetermined and underdetermined problems. Consider our basic equation (3.84)

$$\mathbf{A}\,\mathbf{f} = \mathbf{g} \tag{3.89}$$

as a system of $n \times n$ linear equations (3.87) for n unknown parameters $\mathbf{f}$. Thus $\mathbf{A}$ is a $n \times n$ *square matrix*, assumed regular. The solution is uniquely determined and given by (3.85),

$$\mathbf{f} = \mathbf{A}^{-1}\mathbf{g} \quad . \tag{3.90}$$

This time we have a well–posed problem since the inverse of a regular square matrix $\mathbf{A}$ exists and is unique.

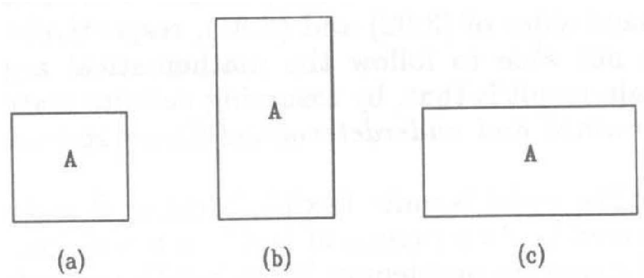


Figure 3.26: The matrix **A** for a well–determined (a), overdetermined (b) and underdetermined (c) system of linear equations.

In the well–determined case of Fig. 3.26(a), the matrix inverse $\mathbf{A}^{-1}$ is regular and uniquely defined. Here the inverse problem is well–posed.

In the other cases (b) and (c), the solution may also be written in the form (3.90) but now $\mathbf{A}^{-1}$ is not uniquely defined. *Both the overdetermined and the underdetermined inverse problems are ill–posed.*

In the *overdetermined* case (b), the system of equations (3.89) is in general inconsistent. It is treated like the adjustment problem in sec. 2.6. A small vector $\mathbf{v}$ is added to $\mathbf{g}$, and the equation is solved subject to the minimum condition

$$\mathbf{v}^T \mathbf{P}\,\mathbf{v} \Longrightarrow \text{minimum} \quad , \tag{3.91}$$

generalizing (2.39) on p. 65. $\mathbf{P}$ is an arbitrary symmetric positive–definite matrix. (If you do not know what positive–definite means, forget it immediately, unless you like high–sounding and pompous words.) The solution is

$$\mathbf{f} = (\mathbf{A}^T \mathbf{P}\,\mathbf{A})^{-1} \mathbf{A}^T \mathbf{P}\,\mathbf{g} \quad . \tag{3.92}$$

In the underdetermined case the system (3.89) is consistent but has infinitely many solutions. Using the minimum condition

$$\mathbf{f}^T \mathbf{Q}^{-1} \mathbf{f} \Longrightarrow \text{minimum} \quad , \tag{3.93}$$

the solution is

$$\mathbf{f} = \mathbf{Q}\,\mathbf{A}^T (\mathbf{A}\,\mathbf{Q}\,\mathbf{A}^T)^{-1} \mathbf{g} \tag{3.94}$$

quite similar to (3.92).

The free choice of the matrices $\mathbf{P}$ and $\mathbf{Q}$ expresses the fact that the solution is not unique. Once $\mathbf{P}$ and $\mathbf{Q}$ have been fixed, however, the solution is unique.

The generalized inverse $\mathbf{A}^{-1}$ is the product of matrices preceding $\mathbf{g}$ at the right–hand sides of (3.92) and (3.94), respectively.

If you are not able to follow the mathematical argument, don't worry. The main result is that, by assuming definite matrices $\mathbf{P}$ and $\mathbf{Q}$, *even overdetermined and underdetermined linear systems get a unique solution*.

Obviously this model is quite flexible because $\mathbf{P}$ and $\mathbf{Q}$ can be chosen rather arbitrarily. As a matter of fact it is not claimed that the underdetermined quantum problem or Whitehead's overdetermined problem should be treated in this way. It is always good to know, however, that well–defined models for the solution of overdetermined and underdetermined systems exist and can be used if necessity arises.

The present considerations also show the basic role of *probability* for the solution of ill–posed inverse problems. Since a unique solution does not exist, one tries to find the "best" solution on the base of statistical considerations. Both least–squares principles (3.91) and (3.93) have a statistical background: the "weight matrix" $\mathbf{P}$ and the "covariance matrix" **Q** represent, so to speak, our statistical information on the problem under consideration.

The importance of additional information. We have just considered statistical information. Most inverse problems are problems of applied rather than pure mathematics, so the physical etc. structure of the problem furnishes fundamental additional information. For instance, geological information may make geophysical (gravimetric, seismic) inverse problems better determined; cf. the articles by G. Anger and H. Moritz in (Anger et al. 1993).

The set of possible solutions. Letting the matrix **Q** vary in (3.94), we may get something like a set of possible solutions of eq. (3.89). The advantage of this way of looking at the solution of an underdetermined problem is obvious: we first get a complete set of possible solutions, from which we may then select a suitable solution by imposing additional conditions.

3.9 Induction, verification, and falsification

To ask whether inductive procedures are rational is like asking whether the law is legal.

Jonathan Cohen

Induction simply does not exist.

Sir Karl Popper

Induction

Frequently, induction is considered the inverse of deduction. *Deduction* proceeds from the general to the particular, using a general law to compute particular observable quantities which then may be compared with actual observations. *Induction* is said to proceed from the particular to the general, using particular observed data to derive the general law. Thus induction is the inverse problem, in the sense of sec. 3.8, of deduction.

This definition of induction is acceptable if we keep in mind that it has a logical status completely different from that of deduction. *Deduction* is a precisely and uniquely defined, straightforward logical process which can be formalized in terms of symbolic logic (sec. 2.1) and may well be performed by a computer. It is a *well–posed problem* in the sense of sec. 3.8.

In the same sense, *induction is an ill–posed problem.* It is impossible to uniquely derive the general law from the data. When Sherlock Holmes claimed that he detected the criminal by rigorous logical deduction from the data — the traces of the crime — don't believe him! He had to solve a problem of induction which in general is far from having a unique solution: it may even have no solution at all, as police statistics show.

The task of the scientist, as "detective of nature", is similarly a problem of induction, which is a difficult "inverse problem" with several solutions (a problem that can be solved by classical mechanics may as well be solved, in a more complicated way, by relativity or quantum theory – but not vice versa) or possibly with no solution at all (finding a "theory of everything", cf. sec. 6.6).

The problem of induction has been one of the most famous and most difficult problems of philosophy, from David Hume (1711–1776) to the present day. Let us start with some simple examples.

(1) *Succession of day and night.* This has been observed since mankind came into existence, and there was never a single exception (see also sec. 3.8). Can we conclude that tomorrow the sun will shine again — at least above the clouds? Pragmatically we all believe that there will be another day, but this cannot be proved logically. *Induction is not a purely logical problem.* If logical procedures such as deduction are called analytic, *induction is not analytic.* It is a physical problem: there will be no tomorrow if the Earth or the Sun explode during the night, or if the Earth has been destroyed by the impact of a huge meteorite. But still we may consider that with high probability there will be another day.

(2) *All swans are white.* Let us assume, for the sake of argument, that, so far, only white swans have been observed. Can we say (a) that *the next* observed swan will also be white and (b) that *all* swans are white? Obviously we can expect event (a) to occur with much higher probability than the general law (b) to be true. Even if zoology claimed that all swans are white (which it does not), a black swan could still

occur: a student might have painted the swan black in order to fool his professor.

It is sometimes said that induction works if there is a certain *uniformity of nature*. This certainly applies to Example 1: the laws of Earth rotation guarantee the succession of day and night if there is no perturbation by a collision with a large meteorite or by an explosion as mentioned above. But will these laws also hold tomorrow?

(3) *Russell's chicken*. We quote from (Russell 1912, Chapter VI): "Domestic animals expect food when they see the person who usually feeds them ... The man who has fed the chicken every day throughout its life at last wrings its neck instead, showing that more refined views as to the uniformity of nature would have been useful to the chicken."

However, expectations govern our daily life as well as science. When I come home and see and smell a nice meal, knowing that my wife is a good cook, I expect the food to be good and healthy. When a physics professor prepares a particularly showy experiment, knowing that he is a skilful experimenter and trusting the laws of physics, he expects the show to be successful. When a botanist plants the bulb of a tulip she expects that, under normal circumstances, she will get the appropriate flower. If a pianist starts a concert with Beethoven's Appassionata and touches the keys accordingly, he and his listeners expect to hear the magnificent sounds of the sonata.

Still, one can never be sure: some jealous colleague may have put the piano out of order right before the concert, etc.

But usually our, often unconscious, trust in the uniformity of nature or its lawful behavior is justified, especially if we use experience and common sense. If we discover a certain regularity and we expect an underlying general law, then we may be justified to assume that the regularity will persist in the near future, and if the regularity continues, we may reasonably assume that we have discovered the corresponding general law. Every new experiment will confirm the law and increase its probability. This may be called *induction by analogy*.

Bayes' theorem. Let H denote a hypothesis, and $p(H)$ its initial "*a priori*" probability, or *prior probability*. Let E denote an event that confirms the hypothesis. Then

$$p(H|E) = \frac{p(E|H)}{p(E)}\, p(H) = \alpha\, p(H) \quad . \tag{3.95}$$

This is the simplest form of *Bayes' theorem*, cf. (Cohen 1989, p. 68). It gives the "*a posteriori*" probability, or *posterior probability*, of H

after the confirming event E. Here $p(E)$ denotes the probability of E in general, and $p(E|H)$ the probability of E if hypothesis H is assumed.

As an example, following Cohen, take H the hypothesis that it will rain within an hour. Let E be the presence of dark clouds. Then $p(E)$ is the probability that dark clouds are present, and $p(E|H)$ the probability of dark clouds if it will rain within the next hour (if hypothesis H is true). Clearly, the probability $p(E|H)$ will be considerably greater than $p(E)$. In fact, $p(E)$ is the probability of dark clouds in general (just looking out of the window without knowing beforehand how the weather is: there may be sunshine or rain). Hence $p(E|H) > p(E)$ and

$$\alpha = \frac{p(E|H)}{p(E)} > 1 \tag{3.96}$$

and

$$p(H|E) > p(H) \quad . \tag{3.97}$$

Thus the "*a posteriori probability*" of H if E is observed is greater than the original ("*a priori*") *probability* of H. The observation of E has increased its probability.

Put $E = E_1$, the confirming event, and assume $p_1(H) = p(H|E_1)$ as the new "a priori" probability. Take a new piece of evidence E_2, for instance, that the barometer has fallen considerably, and compute $p_2(H) = p_1(H|E_2)$ which can be expected to be greater than $p_1(H)$. Consider a third piece of evidence E_3, for instance, the broadcast announcement that very probably it is going to rain within an hour. Use E_3 to improve $p_2(H)$, obtaining $p_3(H) = p_2(H|E_3)$, which by now is presumably already very close to 1.

The mathematical foundation of Bayes' formula is the simple relation

$$p(H|E)p(E) = p(E|H)p(H) \quad , \tag{3.98}$$

well known as one of the most basic formulas in the theory of probability.

Bayes' method can also be generalized to permit a decision between several competing hypotheses, cf. (Jeffreys 1973, p. 31).

Induction and Bayesian inference are particularly useful to find a simple "law" from observations. This "law" may simply be a prescribed mathematical formula such as a polynomial

$$y = a_0 + a_1 x + a_2 x^2 \tag{3.99}$$

whose parameters a_0, a_1, and a_2 are to be estimated by observation ("curve fitting"). Here Bayesian estimation has become a successful competitor to least–squares estimation (sec. 2.6). This is the case in geophysics (Jeffreys 1961, 1973), but also in geodesy which has been a stronghold of least–squares estimation since Gauss, cf. Karl–Rudolf Koch, "*Bayesian Inference with Geodetic Applications*", Springer, Berlin 1990.

An objection against Bayesian methods is that it requires the introduction of prior information, in particular, a priori probabilities. However, even the use of vague prior information leads in many cases to acceptable results.

Conclusion. Karl Popper, known for his pointed statements, said (Miller 1985, p. 104): "Induction simply does not exist, and the opposite view is a straightforward mistake." The theory of inverse and ill–posed problems, however, permits to "tidy up" the problem of induction in the sense of (Jeffreys 1961, p. 8).

Traditional mathematical problems were well–posed. In 1902 the French mathematician J. Hadamard explicitly defined well–posed problems by means of the three conditions given in sec. 3.8: existence, uniqueness, and stability. During the last decades, however, it turned out that many meaningful problems of mathematics and mathematical physics were *ill–posed* or *improperly posed*, for instance, inverse problems, of which some examples are given in sec. 3.8 and more will be found in (Anger et al. 1993), and deterministic chaos which was treated in sec. 3.2. In fact, such ill–posed problems not only have lost the negative connotation implied by their name and have become mathematically respectable: because of their very difficulty and their practical relevance they are currently even quite fashionable in mathematics.

If induction is to be considered a "logically well–posed problem" such as logical deduction of which a unique solution always exists, then Popper is right. The solution of the problem of induction is usually non–unique. Also, it is not a purely logical problem: a priori statistical, physical, etc. information will make the problem of induction more definite. This is exactly analogous to mathematical inverse problems (see end of sec. 3.8). Just as mathematical "ill–posed" problems are mathematically respectable, fascinating, and practically important, in the same way induction is logically respectable, fascinating, and practically important, cf. (Cohen 1989).

(This relation to ill–posed problems is the reason why induction is treated here rather than in Chapter 2 where it might seem to have been better placed, but induction is not a purely logical problem!)

Verification and falsification

The general theories of science have not been found by induction. They are "intelligent guesses" based, of course, on observational data but primarily on other criteria such as logical simplicity and elegance, the theoretician's experience with older theories and, to put it bluntly, by genius and good luck, helped by the "unreasonable effectiveness of mathematics" to be mentioned below. Many people have admired Linné's classification of zoology and botany, but it required the genius of Charles Darwin to discover, on its basis, the *theory of evolution*. Many people have looked on world maps and noted the surprising general similarity of the coast lines of South America and Africa, but it required Alfred Wegener (1880–1930) (according to W. Schröder, preceded by H. Wettstein in 1880) to discover *continental drift* which then became *plate tectonics* as the basic contemporary "paradigm" for geology, geodesy, and geophysics (sec. 3.10). It was the genius of Isaac Newton to discover the general *laws of mechanics* underlying Galilei's principle of inertia and Kepler's laws of planetary motion. The Lorentz transformation was known to Hendrik Lorentz and Henri Poincaré before Einstein recognized its general physical significance, founding the special theory of relativity. Just by a logical extension of the principle of relativity to general coordinate transformations, postulating "general covariance" of the physical laws, Einstein discovered the general theory of relativity, having, as the only empirical basis, the striking equality between "inertial mass" and "gravitational mass" verified with extremely high precision by Roland Eötvös.

Thus the great scientific theories are invented by genius, developed fully by mathematical deduction, and only then subjected to empirical testing or *verification*. Special relativity and quantum theory have been used for a huge number of applications, and each application has confirmed the theories; there has never been a single case where these theories would have been found wrong. The possible tests of general relativity have been fewer in number, but also here the theory passed with flying colors. The first confirmation of general relativity was by observing a solar eclipse by an expedition of the British Royal Society to South Africa in 1919. It was a truly ceremonial occasion when the expedition returned and announced that Einstein's theory was fully confirmed. A dramatic description of this event may be found in (Whitehead 1925, Chapter I).

Sir Karl Popper has maintained that no amount of verification can ensure the validity of a theory, whereas a single *falsification* is sufficient to overthrow it. Therefore, falsification is logically more important than verification. (This was already known to Francis Bacon, 1561–1626.)

This has an important counterpart in mathematics: a single *counterexample* will invalidate the most sophisticated mathematical proof, showing that it must contain a logical error. Counterexamples are also useful to find out the exact scope of a mathematical theorem, and are often used in this way, both by active researchers and by professors to give their students a feeling of the applicability and the limitations of a theorem.

However, this holds for pure mathematics and pure logic. Actual data are almost always affected by uncertainties and errors. Thus even Popper's falsification is not absolute: the falsification may only be *apparent*, caused by a measuring error, whereas in reality, the theory is true. This is not a theoretical speculation: many modern experiments operate in the gray zone between error and reality: it may be difficult to decide whether a certain small effect is "real" or due to measuring errors.

Following Gauss, astronomers, geodesists, and geophysicists have developed a strong feeling for the importance of measuring errors: they have acquired a psychological habit of basic mistrust towards observations. This is not because they do not respect observations. Quite on the contrary: they respect them so much that they want to know exactly how reliable and how accurate they are.

Measurements should always be accompanied by their "*root mean square error*", or "*standard error*" which is a measure of their accuracy, e.g. for a measurement of length l:

$$l = 124.327\,\mathrm{m} \pm 2\,\mathrm{mm} \quad .$$

Another principle is "*Eine Messung ist keine Messung*", one measurement is no measurement unless it has been checked by another measurement.

This is an immediate *practical objection against falsification*: one cannot be sure that the *one* falsifying measurement is really correct.

Thus both verification and falsification are necessary: as soon as a promising new theory has been published, it will immediately be subjected to experiments, if possible of the "crucial" type (see below), and *every attempt of verification is also an attempt of falsification*, depending on the outcome of the experiment. So a theoretician need not worry about verification or falsification of his theory: his experimental friends will try to verify it and his opponents will be most happy to falsify it. The more experiments are performed, the more the effect of measuring errors will be reduced.

It is fundamental that physicists and natural philosophers put much emphasis on Heisenberg's uncertainty principle, but ordinary measuring errors, which occur much more frequently and may be much larger, also deserve their attention; cf. (Jeffreys 1961, pp. 13–14). Jeffreys' books merit particular respect because they were written by a scientist of enormous experience with actual "dirty" data.

Many extreme logical or philosophical conclusions do not apply to our real "fuzzy" world because of their very subtlety: they are used as razors, not for splitting hairs (which would be appropriate) but for cutting trees.

The results of deduction are logically true if the process of deduction is done correctly. The results of induction, including verified or, better, not–yet–falsified, theories, can, at best, be probable or acceptable on a hypothetic basis. No physicist will consider relativity or quantum theory "absolutely true", in the same sense as 2+2=4. He regards them as excellent and unsurpassed working tools, even as *correct*, but only in the sense of exceptionally good approximations; cf. also sec. 6.5. In this way, also Newtonian mechanics remains "correct": for small velocities and phenomena above the quantum level.

Crucial experiments. Nevertheless, there are crucial experiments which really permit to decide between two theories or hypotheses. One of the most famous crucial experiments is the *Michelson–Morley experiment* which shows that the velocity of light is the same along all directions on the moving Earth. This is incompatible with classical mechanics, eq. (3.36) on p. 91, but perfectly compatible with the special theory of relativity.

It is instructive to study this case by means of Bayes' formula. Generalizing (3.95) to the case of two competing hypotheses H_1 (*classical mechanics*) and H_2 (*relativity*), we have

$$p(H_i|C) = \frac{p(H_i)p(C|H_i)}{p(H_1)p(C|H_1) + p(H_2)p(C|H_2)} \tag{3.100}$$

$$= A\,p(H_i)p(C|H_i) \tag{3.101}$$

where A is a constant which is independent of whether H_1 or H_2 is accepted; i denotes 1 or 2.

Now, $p(H_i)$ denotes the *prior probability* of hypothesis H_i, $p(H_i|C)$ its *posterior probability* after performing the crucial experiment C, and

$p(C|H_i)$ is called the *likelihood* which is the probability of C on the basis of H_i. We may thus write

$$\textit{posterior probability is proportional to prior probability} \times \textit{likelihood} \tag{3.102}$$

(Jeffreys 1973, pp. 30–31). It is presupposed that either H_1 or H_2 is true.

Now $p(C|H_1)$, the probability of the result of the Michelson–Morley experiment (the light velocity on Earth is the same in all directions) on the basis of classical mechanics is very small; we put

$$p(C|H_1) = \epsilon \quad . \tag{3.103}$$

If the Michelson–Morley experiment would be absolutely true, then $\epsilon = 0$, but no experiment is absolute so we take ϵ to be a small number > 0. The more reliable the crucial experiment, the smaller is ϵ.

Similarly

$$p(C|H_2) = 1 - \delta \tag{3.104}$$

since on the basis of H_2 (relativity), C theoretically must be true, but we take $\delta > 0$ but small for the same reason as $\epsilon > 0$.

Let us first assume that the prior probabilities of H_1 and H_2 are equal:

$$p(H_1) = p(H_2) = \frac{1}{2} \quad . \tag{3.105}$$

At any rate

$$p(H_1) + p(H_2) = 1 \tag{3.106}$$

since either H_1 or H_2 are considered true.

Then (3.100) gives

$$p(H_1|C) = \frac{0.5\epsilon}{0.5\epsilon + 0.5(1-\delta)} = \frac{\epsilon}{1+\epsilon-\delta} \doteq \epsilon \ , \tag{3.107}$$

$$p(H_2|C) = \frac{0.5(1-\delta)}{0.5\epsilon + 0.5(1-\delta)} = \frac{1-\delta}{1+\epsilon-\delta} \doteq 1-\epsilon \ , \tag{3.108}$$

stating that H_1 (classical mechanics) has become very improbable and that H_2 (relativity) has been confirmed, the better, the more reliable the crucial experiment is.

Let us now assume that the prior probability of H_1 is high:

$$p(H_1) = 0.9 \quad , \tag{3.109}$$

so that

$$p(H_2) = 0.1 \tag{3.110}$$

in view of (3.106). Then (3.100) gives

$$p(H_1|C) = \frac{0.9\epsilon}{0.9\epsilon + 0.1(1-\delta)} = \frac{9\epsilon}{1-\delta+9\epsilon} \doteq 9\epsilon \ , \tag{3.111}$$

$$p(H_2|C) = \frac{0.1(1-\delta)}{0.9\epsilon + 0.1(1-\delta)} = \frac{1-\delta}{1-\delta+9\epsilon} \doteq 1 - 9\epsilon \ . \tag{3.112}$$

Thus, the crucial experiment has made to drop the probability of classical mechanics from 0.9 to 9ϵ. If $\epsilon = 10^{-6}$ (the experiment is very reliable), then

$$p(H_1) \doteq 10^{-5} \quad , \tag{3.113}$$

$$p(H_2) \doteq 1 - 10^{-5} \quad , \tag{3.114}$$

which shows that relativity is confirmed also in this case: $p(H_2)$ has risen from 0.1 to almost 1!

In fact, the assumption $p(H_2) = 0.1$ is not too high because, even before the Michelson–Morley experiment it was known that the Lorentz transformation (3.38) on p. 92 holds for electrodynamics: Maxwell's equations are invariant with respect to (3.38); this was precisely what Lorentz showed. So in view of the universality of physics, there was a certain a priori probability that H_2 would hold also for mechanics.

We see that basically the same result is obtained for very different prior probabilities. This indicates that the choice of the prior probabilities is not very essential.

We shall have much more to say on the laws of nature in following sections. There is a long way from "Russell's chicken" to the "eternal inexorable laws of nature" of romantic poets and happy–minded scientists, and a still longer way to "Schrödinger's cat" ...

The unreasonable effectiveness of mathematics. Already at this point, however, we mention, for the first time in this book, a fact that has intrigued physicists from Kepler to Einstein. The great quantum physicist Eugene Wigner has called it the "unreasonable effectiveness of mathematics in the natural sciences". Penrose (1989, p. 430) puts it as follows ("SUPERB" theories are, e.g., special and general relativity and quantum mechanics, cf. p. 260):

It is hard for me to believe, as some have tried to maintain, that such SUPERB theories could have arisen merely by some random natural selection of ideas leaving only the good ones as survivors. The good ones are simply much *too* good simply to be the survivors of ideas that have arisen in that random way. There must, instead, be some deep underlying reason for the accord between mathematics and physics, i.e. between Plato's world and the physical world.

Summary. Induction does exist. In simple cases, induction by analogy can be used with some care: Bayes' theorem may help, also in the estimation of parameters for simple laws given by a function containing several parameters. More general laws such as relativity require the creative mind of a great scientist; verification or falsification by experiment are necessary but can be expected almost automatically to be performed by the scientific community. Crucial experiments are especially important.

Additional reading. A nice introduction is the chapter on induction in (Russell 1912). If you read what Russell (1948), Carnap (1950, 1966), Popper (1977; Miller 1985), and Jeffreys (1961, 1973: if you don't need them, disregard the formulas but do read the text) have to say about induction, and if you conclude with (Cohen 1989), then you should know almost all the relevant present views on the topic.

3.10 The structure of scientific revolutions according to Kuhn

Led by a new paradigm, scientists
adopt new instruments and
look in new places.

Thomas Kuhn

T.S. Kuhn (1970) has given a theory of the history of scientific revolutions which has attracted general attention (but also some controversy as usual in such cases). Scientific revolutions introduce not only new theories but imply a change in the general scientific climate. Old *paradigms* are replaced by new ones. A paradigm is more than a scientific theory: it is a way of thinking, a way of looking (Greek: *theoria*) at nature.

Perhaps the most famous new paradigm was the *Copernican revolution*. The Earth was no longer the center of the universe. It is one of the planets that orbit around the Sun. Kepler's laws of planetary motion and the mechanics of Newton, Legendre, and Laplace were the consequences.

Another revolutionary change of paradigm was *biological evolution*. The static system of botanical and zoological classification of Carl von Linné (1707–1778) was made into a dynamic theory of evolution by Jean Lamarck (1744–1829) and Charles Darwin (1809–1882).

Immanuel Kant (1724–1804) performed a "Copernican revolution" in philosophy, emphasizing the role of the subject. In this way, he founded the great school of *German idealism* (Fichte, Schelling, Hegel), but he also influenced the contemporary philosophy of science.

The logical discoveries of George Boole (1815–1864), Gottlob Frege (1848–1925), Giuseppe Peano (1858–1932) and Bertrand Russell (1872–1970) lead to analytical philosophy and the modern *theory of science*, or philosophy of science.

In physics, of course, we have the new paradigms of *relativity* and *quantum theory*, and most recently, *chaos theory*.

The recent development in biochemistry and molecular biology started modern *genetics. Reductionism* (life can be reduced to chemistry and physics) is not a logical consequence of these developments, but is rather generally accepted at least as a working hypothesis.

Cybernetics, *system theory*, *catastrophe theory*, *complexity theory*, and *synergetics* also have influenced modern scientific thinking, closely related to the advent of electronic computers. These concepts will be explained later (sec. 4.2).

In geosciences, we now have the paradigm of *plate tectonics*. The Earth's surface consists of a number of continental plates which move with a speed on the order of 5 cm per year. Colliding with each other, they pile up mountain chains such as the Rocky Mountains and the Andes, but also the Himalayas and the Alps, accompanied by earthquakes and volcanism. It started with Alfred Wegener's continental drift (sec. 3.9) published in 1915, but it was generally recognized only in the sixties.

A scientific revolution does not occur only because what Kuhn calls "normal science" has been proved wrong or, in Popper's terminology, "falsified". The old epicycle theory of planetary motion by Ptolemy (2nd century A.D.) could have easily been adapted to the increasing measuring accuracy by adding one more epicycle or two; it simply grew too complicated to retain credibility. Kepler was motivated not only by his belief in the simplicity and harmony of the world (which, in some way or other, is shared by modern scientists as well) but also by mystical speculations: he related the planetary orbits to the five Platonic solids (p. 118). Genius is inexplicable: reasonings are frequently quite irrational and even the significance of the results is not always properly understood by their discoverers.

Lorentz found the equations of the "Lorentz transformation", but only Einstein and Minkowski realized their revolutionary physical significance: the special theory of relativity was created. Einstein tried to generalize the invariance with respect to Lorentz transformations to "general covariance": the general theory of relativity was discovered. Now, however, we put more emphasis on curved space–time than on the democratic equivalence of all reference systems which general covariance means (sec. 3.4).

In fact, general relativity is not the only space–time theory of gravitation; there are other theories that explain the phenomena equally well. Still, few physicists doubt that general relativity, in a way, is the "best" theory of gravitation, in view of its incomparable internal perfection. (Remember the "unreasonable effectiveness of mathematics" of sec. 3.9!)

Still, the great French mathematician Henri Poincaré (1854–1912) believed that Euclidean geometry would never be given up precisely because of its internal perfection. In fact, gravitation can be handled (and is handled practically) in many cases by a basic Euclidean geometry (and Newtonian physics) plus "relativistic corrections", but the conceptual framework of general relativity is still recognized superior to this way of putting gravitational "epicycles" onto Euclidean geometry.

Thus in many cases, scientific revolutions are not always "logically necessary": normal theory could still continue for quite a while, perhaps by piling up "epicycles", but finally, the edifice collapses and is replaced by a new "paradigm".

Like many revolutions, new paradigms tend to be dogmatic. Moser (1989, p. 145) has pointed out that a molecular biologist would commit scientific suicide if he would support some kind of vitalism (sec. 4.5) instead of the current dogma of reductionism. A geophysicist who flatly contradicts plate tectonics has little chance to get his work published in a reviewed journal.

Physicists, especially renowned ones, are relatively free to express also views that run counter to the prevailing opinion. This attitude has perhaps been fostered *in physics* by people like Niels Bohr who said that "a theory must be very crazy to be true". In fact, mathematical models that run so much counter to common sense as Everett's "many–world interpretation" of quantum theory or the theories of supersymmetry and superstrings are being seriously discussed in physics. This is what makes the extravagant mathematical models and the incredibly imaginative thinking of modern physicists so attractive and important philosophically: they show how far logical speculation can go. The most imaginative and unorthodox great contemporary physicist is probably John A. Wheeler; the book (Wheeler 1994) thus is enormously important regarding possible philosophical implications of modern physics, although it is not quite easy (Rabbi type 2 to 3).

This also seems the reason why great physicists rarely content themselves with the current school of analytical philosophy. Einstein was a pantheist of Spinoza type, Gödel was an objective idealist who rejected the reality of space–time, Schrödinger's views were close to Buddhism, Bohr was a natural dialectic thinker, Weizsäcker likes classical philosophy as much as physics ...

The contemporary science business and paper industry with its refereed journals is probably necessary. Still, one wonders how Max Planck could have published his first paper on quantum theory in a refereed journal ... Nevertheless, even today, novel ideas get their chance, e.g., through presentation in scientific meetings.

The scientific climate prevailing today is very well characterized by the following anecdote kindly communicated to the author by Professor Paul Melchior from his personal collection:

Why God Failed to Receive University Tenure

- He had only *one* major publication.
- It was not published in a refereed journal.
- It had no bibliography.
- It was not in English.
- He did not even entirely write it Himself, but had a number of mostly anonymous co–authors.
- It may be true that He created the world, but what has He done or published since?
- His relations with His scientific colleagues are often strained.
- The scientific community is having a very rough time trying to rediscover His unpublished results.

Chapter 4

Systems, information and evolution

4.1 Feedback, regulation, and downward causation

A variation of one–half degree centigrade in the body temperature is generally a sign of illness, and a permanent variation of five degrees is scarcely consistent with life.

Norbert Wiener

In sec. 1.1 we have seen that, in man and animals, the hypothalamus serves, among other functions, as a thermostat regulating body temperature in order to maintain this temperature at a constant level within very narrow bounds. Feedback and regulation mechanism play a fundamental though largely unconscious role in the working of the human body, in movements such as walking, and may even be responsible for what we call the action of mind on our body, such as reaching for a book which we need when working at a problem. Such an action from a higher level (thinking or willing) to a lower level (bodily movement) is called *downward causation*. Also the activity of a computer when we feed in a program and data, or instruct it to display a certain information on the screen, are examples of downward causation (causing activity on the "lower" hardware level by providing input on the "higher" software level).

An essential concept is *feedback*. A teacher gives a course, and the students listen with more or less attention. If the teacher cares about being understood, he needs information about how his teaching is received by the students. This is an example of feedback. Feedback may comprise questions put during or after the lecture, criticism provided by the students, or even the average result of the examinations, as well as watching the students' behavior during the lecture, from bright attentive faces to yawning and sleeping.

The direct "feed–forward" of information from teacher to student, together with this feedback, forms a characteristic loop, the *feedback loop* (Fig. 4.1).

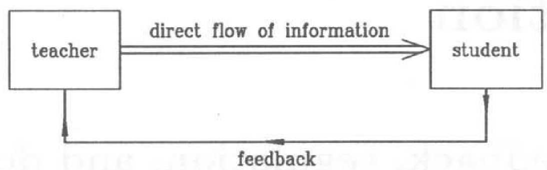


Figure 4.1: A feedback loop

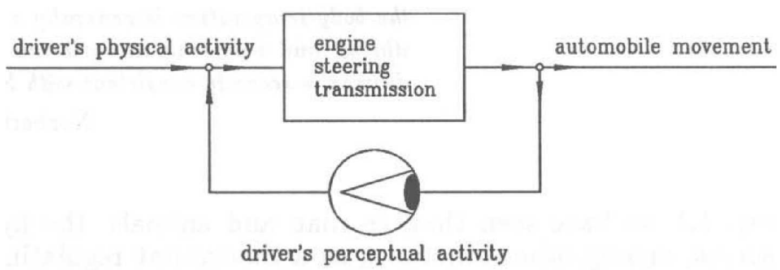


Figure 4.2: The feedback loop for automobile driving

Another example from every–day life: the driving of an automobile (Fig. 4.2). The driver's (physical) activity comprises turning the steering wheel, alternately depressing gas pedal and brake pedal, shifting gears (in cars which are not yet equipped with automatic transmission), etc. The final goal is reaching a certain destination, the immediate goal, however, is staying on the road and avoiding accidents. This goal requires the driver's constant observation of the momentary situation, comparing the car's movement to the movement desired by the driver and correcting deviations by appropriate physical activities such as turning the steering wheel, accelerating or braking. The new course of the automobile is again fed back to the driver, deviations are observed and corrected, etc.

This example also helps introduce two additional concepts. The automobile acts as some kind of *servomechanism* for man, helping him/her to cover large distances much faster than by walking. Servomechanisms in the narrower sense frequently replace human movement and observation of deviations (from a goal) by automatic functions. This happens, for instance, with a stabilized platform: deviations from a fixed orientation are measured by gyroscopic sensors and activate servomotors which restore the desired fixed orientation. Modern airplanes are frequently equipped with autopilots which, using inertial, radar, altimeter, etc. information will relieve the pilot from considerable purely routine work, frequently do certain specific and well–defined tasks more precisely and accurately than the human pilot could do.

A second important concept is the name, *cybernetics* or *kybernetics*, for a whole new discipline: control and communication in man, animals, and machines. This name was introduced in 1947 by Norbert Wiener, cf. (Wiener 1961, p. 11). It is derived from the Greek word *κυβερνήτης* (*kybernetes*), which means steersman or pilot.

It shows that machines exist which are designed to perform a certain task (Aristotle's "*final cause*") through appropriate mechanisms operating according to the usual laws of causality in the physical sense (Aristotle's "*efficient cause*"). Thus the gap between physics (believed to be subject to usual causality) and biology (where "goals", "aims", "purposes" and similar "final causes" seem to govern the behavior of animals and even the course of evolution) has at least partially been bridged. Earlier, a "vital force" has been considered necessary to explain "finalistic" animal behavior, but, according to Wiener (1961, p. 44): "In fact, the whole mechanist–vitalist controversy has been relegated to the limbo of badly posed questions".

Other examples of servomechanisms operated by man are bulldozers and other excavating machines, replacing shovels and similar tools. In fact, such tools and machines can be regarded as extensions of our *body* to perform certain tasks much better than using our hands only.

Now it is basic that our hands, feet and other bodily organs may also be considered instruments or tools ("servomechanisms") directed by our *mind* (or by the neural activity in our brains, if you prefer) to perform certain operations, for instance reaching for a glass of water and bringing it to our mouth. This operation is constantly controlled by visual and other feedback, as we easily recognize if we try to perform it in darkness.

Here we have an instructive case taken from (Wiener 1961, p. 95). A patient tries to perform the task just mentioned. He will take the glass and move it violently in a direction towards his mouth. He moves, however, too far in an initial direction and tries to correct this by an equally violent movement in an opposite direction, until his motion becomes nothing than a futile and violent oscillation. Needless to say, he will have spilled the glass of water before even having it brought close to his mouth. What has happened?

The patient is suffering from what is called cerebellar tremor. The feedback mechanism is abnormally strong, so instead of a steady motion (with small oscillations, to be sure), we have oscillations which grow indefinitely and never "converge" to achieve the desired goal.

Or consider a loudspeaker amplifier. If the amplification becomes too high, the amplitude of the oscillation grows in an uncontrolled way, resulting in a shrill and penetrating uniform noise.

Here we have met with an important concept: *amplification*, which is basic in our context. An amplifier converts the weak voice of a speaker to the majestic message coming from a loudspeaker. The excavating machine enormously amplifies the signal input provided by the operator's hands. The extremely weak electric and/or chemical nerve signal coming from our brain must also be amplified enormously so that we can lift a heavy load, etc.

Thus all our bodily movements are possible only through elaborate processes of feedback and amplification. The muscles contain sensors whose outputs are fed back to the brain. A wonderful example is walking. We usually are not aware how delicate and complicated these processes are until we try to walk at night or move on an icy surface. In fact, it is extremely difficult to construct a walking machine.

As a final simple example consider the working of a *thermostat*. Its function is to keep the temperature of a room constant even if external circumstances (outside temperature etc.) vary. The desired temperature serves as input. A thermometer measures the actual temperature. The difference: desired minus actual temperature, activates the heating or cooling system such as to counteract the change (*negative feedback*). Possible modes of behavior are illustrated by Fig. 4.3. Starting from an initial temperature, a well–regulated thermostat will approach the desired temperature through small oscillations whose amplitude rapidly decreases to zero. This is the usual case of *negative feedback*.

Consider now the right–hand side of Fig. 4.3. The graph of "actual temperature", starting at the same initial temperature, moves in the right direction, towards the "desired temperature", but widely over-

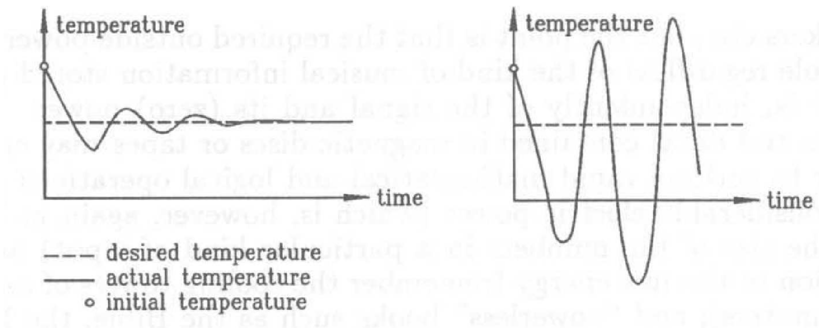


Figure 4.3: Negative and positive feedback

shoots the mark, and then oscillates back and forth with rapidly increasing amplitude, so that a person in the room would not even have the choice between being heated and frozen to death: this would happen in rapid alternation. This corresponds to *positive feedback* which clearly does not provide the desired stabilization of temperature, quite similar to the cerebellar tremor mentioned above.

Mathematically, this is evident: negative feedback corresponds to exponentially decaying cosine (or sine) functions

$$e^{-\alpha t} \cos \omega t \ , \quad (\alpha > 0) \tag{4.1}$$

whereas positive feedback, expressed by

$$e^{\alpha t} \cos \omega t \ , \quad (\alpha > 0) \tag{4.2}$$

gives exponential growth of the oscillation. It is clear that the behavior of the above–mentioned patient suffering from cerebellar tremor, as well as the shrill loudspeaker noise, are described by functions of type (4.2).

Negative feedback also provides a stabilizing element in amplification, as every electronic engineer and every hi–fi fan knows.

Downward causation. In principle, an arbitrarily small signal can be amplified with high fidelity to an almost arbitrarily high power. But what about if the input signal has *no power at all*? Examples are "powerless" compact discs which reproduce all the original magnificent volume of sound of the Alpensymphonie of Richard Strauss. Of course, a compact disc player needs considerable outside power to activate the loudspeakers etc., but the point is that the required outside power must be available regardless of the kind of musical information stored in the disc, that is, independently of the signal and its (zero) power. Input (programs and data) contained in magnetic discs or tapes may cause a computer to perform rapid mathematical and logical operations which require considerable electric power (which is, however, again independent of the size of the numbers in a particular kind of input) and its degradation to thermic energy (remember the cooling towers of modern supercomputers); and "powerless" books such as the Bible, the Koran or, to a lesser extent, the works of Marx, Engels, and Lenin, have completely changed the face of the Earth.

This is possible through the use of some "reading equipment" such as laser or magnetic scanners, or people reading those books, which convert the "written" input into small electrical signals (including the eye converting the optical information into nerve impulses!), which are then greatly amplified and produce the results just described.

Electronic computation, hardware being activated by software, is a beautiful example of *downward causation*: information on a higher level (software) causes action on a lower level (hardware). Note that *this hardware motion is fully governed by physical laws*, in particular the laws of electronics! "Software laws" activate the appropriate "hardware" (physical) laws; "final causes" act through activating the appropriate "efficient causes". Software provides *initial conditions* causing the computer to start working, as well as *boundary conditions* which regulate its work during the computation. Generally, the term "boundary conditions" is used as a compact synonym for both initial conditions and boundary conditions in the narrower sense.

The interaction of mind with matter may well be of this kind. Objections that mind does not possess physical energy and cannot therefore act on matter, lose their force if we compare the action of mind on matter to the action of software on hardware in a computer. For more details cf. (Popper and Eccles 1977, Chapter E7), (Eccles 1994), (Globus et al. 1976), and (Margenau 1984). Haken (1981, p. 196) discusses interaction at various levels.

A thorough discussion of mind–matter interaction can only be given by quantum theory since quantum effects are believed to play a role in neural activity. A comprehensive reference is (Stapp 1993); also (Lockwood 1989) and (Penrose 1989) provide valuable insight. One now even speaks of "quantum computers"!

Finalism in physics. Even in classical physics (mechanics and optics) there are principles that seem to express some purpose and are thus "final causes" in the sense of Aristotle. E.g., we have *Fermat's principle*: light moves from point A to point B along a path for which the travel time is a minimum, or the *principle of Euler–Maupertuis* which states that a mechanical system moves in such a way that a certain integral quantity, the "action", is minimized (principle of least action), cf. eq. (3.20) in sec. 3.1, p. 78.

As we have already mentioned there, these "finalistic" principles can, by a mathematical procedure called the calculus of variations, be transformed into the ordinary differential equations in which classical mechanics usually is formulated and which are *the* expressions of classical causality. Again, "final causes" automatically produce the required "efficient causes"!

The principle of least action acquires a particular interest in the light of Feynman's formulation of quantum mechanics in terms of "path integrals". Such integrals are sums (integrals) over all possible paths leading from A to B. The smaller the Planck constant h (sec. 3.5) is, the more influence is concentrated at paths very close to the "classical path" of least action: the effect of paths far from the classical path almost vanishes because of destructive interference, cf. (Misner et al. 1973, p. 499). For the classical case, $h \to 0$, only the path of least action "survives".

Does not this remind you of a "survival of the fittest" in the sense of Darwin's evolution? And is it too c extravagant to regard Darwin's principle (survival of the *fittest*) as an optimum principle which is some kind of biological analogue to the finalistic principles of Fermat, Euler, and Maupertuis in physics?

And to continue these pretty crazy ideas, is the synergetic cooperation in the sense of Maturana and Varela (1987) and Jantsch (1980), complementary to Darwin's "law of the jungle", a biological counterpart to the Gaussian principle (3.13) on p. 75 which "adjusts" Newton's law to a "different physical environment" of a curved surface, accepting the environment and "cooperating with it"? Cooperation and "democratic adjustment" (see end of sec. 2.6) rather than struggle for survival?

Anyway, finalism or "teleology" seems to act quite obviously in biological systems, as already recognized by Aristotle. The present considerations (from thermostats to downward causation and least–action principles in physics) may thus help bridge an important gap between living and nonliving systems, perhaps even between mind and matter.

4.2 Self–organization

Self–constructibility is an emergent property of a complex system.

Sir Alan Cottrell

Self–organization versus external organization. Beethoven's Fifth Symphony is being performed. The orchestra is playing, precisely following the instructions of the conductor and realizing every nuance of his conducting. This performance of Beethoven's symphony is an example of *external organization* done by the conductor.

Four players perform a string quartet, perhaps also by Beethoven. No conductor is in view, the four players are on more or less the same level, and their equally precise performance is achieved through skilful and sensitive cooperation. This clearly is an example of *self–organization* by cooperation or "synergy", to use a term, *synergetics*, which is becoming almost as fashionable as "cybernetics" introduced in the preceding section.

The distinction between external organization and self–organization is not so absolute as it might look at first sight. The conductor also listens to the orchestra and tries to lead it in a cooperative way, making use of its strengths, taking into account its weaknesses and even attempting to compensate mistakes which the musicians may have made. The orchestra players not only watch the conductor (although some seem to pay no attention to him at all) but also listen to the play of their fellow musicians. This may result in a performance of overwhelming delicate precision and passion.

On the other hand, in the string quartet, the first violinist, *primus inter pares*, nevertheless plays a more or less leading role.

Staying with musical examples, there are orchestras without conductor which may be very good (though not so good as excellent orchestras under a brilliant conductor). The organist leading the congregation's singing in church is not a dictator either, he subtly adapts to the singing of the people, just as a good conductor will not conduct "against" his orchestra. And a pianist playing against a singer, rather than accompanying her, will soon have lost his job.

Other examples. Already in the 17th century, Huygens noticed that two pendulum clocks hanging on the same wall tend to synchronize, and he suspected correctly that this phenomenon of *entrainment* was caused by a coupling of the two clocks through the elasticity of the wall.

The synchronized playing of the two hands on the piano is mainly controlled "externally" through the player's ears and brain, but the coordination of two players performing a musical piece for four hands, is clearly another form of self–organization.

Self–coordination and "cooperation" of light waves with electrons in a ruby laser, generating almost perfectly monochromatic light, are a favorite example of the founder of "synergetics" (Haken 1981, Chapter 5).

Waves and convection cells. I have always wondered about ocean waves. It is clear that they are driven by wind, but why are there waves, and if there are waves, why are their crests where they are, and not a couple of centimeters away? (See Fig. 4.4.)

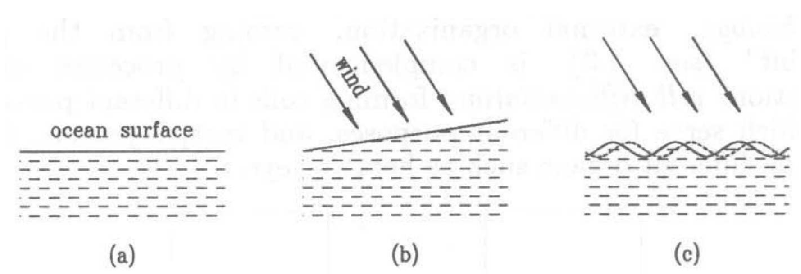


Figure 4.4: Water waves (c) formed by "self–organization" from an unstable inclined plane surface (b)

A uniform wind will first slightly tilt the plane ocean surface (Fig. 4.4 (b)). This tilted plane is *unstable* and hence breaks down. Then the wave motion, which is a stable motion determined by parameters of wind and water and by atmospheric pressure, takes over. What remains undetermined and "random", however, is the position of the wave crests (the "phase"): the same factors might as well have formed a wave translated to the right (as shown by the broken line) or to the left. But once the symmetry of the smooth plane surface has been broken, the phase is determined.

Another well–known phenomenon are *convection cells* formed by heating the lower surface of a fluid layer (Bénard problem), cf. Fig. 4.5.

The latter is clearly relevant for convection in the Earth's viscous mantle, causing the movement of continental plates in plate tectonics.

Also regular *rolling cloud patterns*, frequently visible in an otherwise clear sky, are caused in a similar way.

Chemistry and biology. Self–organization also occurs in *chemistry*,

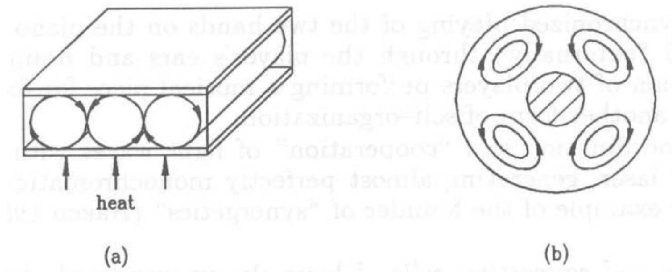


Figure 4.5: Convection (a) in a plane slab and (b) in a spherical shell

forming spatial and temporal wave patterns (chemical clocks, Belousov–Zhabotinsky reaction, cf. (Haken 1981, p. 75)).

In *biology*, external organization, coming from the genetic "blueprint" (sec. 4.3), is complemented by processes of self–organization: *cell differentiation*, forming cells in different parts of the body which serve for different purposes, and *morphogenesis*, the formation of different organs such as heart or eye.

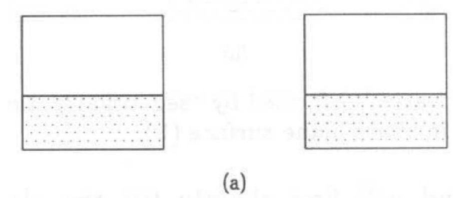


Figure 4.6: Two separated cells with equilibrium concentration of substance A (shaded region)

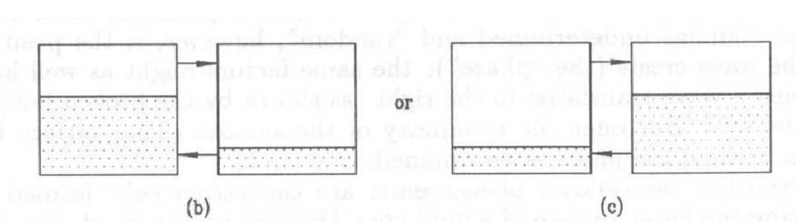


Figure 4.7: Interaction between the two cells leads to two new configurations of equilibrium)

Very little is known in this field, which furthermore is rather controversial. We mention now only one of the simplest models for cell differentiation due to the famous mathematical logician A.M. Turing (Haken 1981, p. 104). Consider two separated identical cells (Fig. 4.6), in which a substance A is contained in a certain equilibrium concentration. Now we admit an exchange of substances between the two cells. This may make the initial concentration of A unstable, and a new equilibrium is reached, in which either cell 1 or cell 2 gets a higher concentration, both possibilities being equally probable (Fig. 4.7).

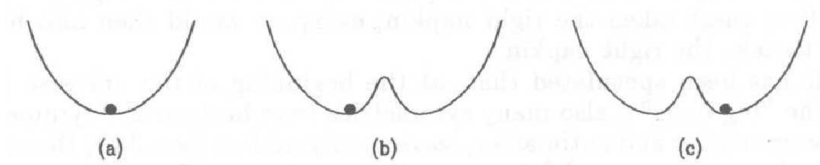


Figure 4.8: Potential energy and symmetry breaking. The cases (a), (b), (c) correspond to Figures 4.6 and 4.7

Equilibrium corresponds to a minimum of potential energy, illustrated by a small sphere permitted to roll without friction on the "potential curves" of Fig. 4.8. The lowest position of "rest" corresponds to minimum potential energy.

Symmetry and symmetry breaking

Symmetric shapes frequently are characterized by minimal potential energy and seem — for this and other reasons — be particularly favorized by nature; cf. the beautiful book (Weyl 1952) and its modern counterpart (Mayer–Kuckuk 1989). Some well–known symmetric shapes are the sphere and the cube, the equilateral triangle, the square, the regular pentagon (5 corners), and the regular hexagon (6 corners) known from honeycombs. Flowers exhibit beautiful symmetries; and *ornaments*, based on symmetry, are as old as human culture.

Raindrops are spherical, *rotating planets* are (nearly) ellipsoidal. *Snowflakes* also show beautifully symmetric patterns. *Crystals* formed by growth in an appropriate solution are "self–organized" (in contrast to a cut diamond formed by the "external" organization designed by man). Symmetric patterns in biology reach from *honeycombs* to the *double helix* of DNA, regarded as an "*aperiodic crystal*" or "aperiodic solid" by Schrödinger (1944); cf. sec. 4.3. Crystals, having minimum energy, are particularly stable, and this fact, according to Schrödinger, may also account for the stability of genes.

Turing's model for cell differentiation involves a particularly simple case of *symmetry breaking*. Cf. Figures 4.6 and 4.8(a) (symmetry) and Figures 4.7 and 4.8(b) and (c) (broken symmetry). Such a symmetry breaking is a very typical and frequently occurring phenomenon.

Think of a completely symmetrically arranged round dinner table. Should I take the napkin to my left or that to my right? I prudently wait until the most courageous dinner guest takes the left napkin. Then the symmetry is broken and everyone will take the left napkin. Had the first guest taken the right napkin, everyone would then also have had to take the right napkin.

It has been speculated that, at the beginning of the universe (after the "big bang"), also many symmetries were broken: the symmetry between matter and antimatter, leaving only matter (sec. 3.6); the symmetry between past and future, creating the "arrow of time" (sec. 3.7); the symmetry between the four basic forces (sec. 3.6), etc.; cf. (Jantsch 1980).

Models for complexity

We have already considered Turing's simple model for cell differentiation (Fig. 4.7). Turing in 1952 also gave the first simple mathematical model for morphogenesis (Mayer–Kuckuk 1989, p. 212).

A very detailed mathematical theory for such processes has been elaborated by the French mathematician René Thom. It has become famous by the name of "*catastrophe theory*" and is highly ingenious but also quite controversial, partly also because of its provocative name (it is really "mathematical morphogenesis").

The last 50 years have seen a number of closely related theories trying to deal with the phenomena of *complexity*, *self–organization* etc. It is typical that all these theories attempt to treat biological systems and inorganic systems, both natural (convection cells) and man–made (machines, computers), by the same methods.

It thus is quite characteristic that Norbert Wiener, the founder of cybernetics, has defined it as "control and communication in the animal and the machine".

Thus we have

– *cybernetics*: Wiener (1961), cf. sec. 4.1;

– *catastrophe theory*: Thom (1975);

– *chaos theory*: sec. 3.2; cf. (Abraham and Shaw 1992; Lorenz 1993);

– *synergetics*: Haken (1981);

– *complexity theory*: (Waldrop 1992), (Lewin 1992), (Gell–Mann 1994).

Ilya Prigogine ("non–linear thermodynamics", "dissipative structures") and Manfred Eigen ("hypercycles") did not invent such catchy names, but they at least got the Nobel Prize, cf. sec. 4.3.

Some characteristic features of complexity. It is very difficult, if not impossible, to define complexity, because it comprises many features and is a field without definite boundaries. Let us nevertheless try to list some features of complex systems.

– A *great number of elements* seems to be necessary but by no means sufficient. A heap of sand contains many grains, but this does not yet make it a complex system. A biological organism consists of many cells and is a prototype of a complex system because:

– a complex system possesses a rich *structure*, an order which is intrinsic rather than imposed from the outside, e.g., an animal versus an automobile. Both kinds of order are, however, complementary: the genetic structure provides a certain basic information "from outside", which is to be supplemented by cell differentiation (e.g., a blood cell vs. a liver cell) and morphogenesis through *self–organization*, cf. sec. 4.3. An adaptation to the environment is frequently implied; Gell–Mann (1994) speaks of *complex adaptive systems*.

– The intrinsic order of a complex system is *dynamic* rather than static: it must always defend itself against *chaos.* Think of a warm–blooded animal: it must permanently strive to keep its bodily temperature constant, in spite of the usually colder environment with all its random temperature changes. Another example is a person who must constantly endeavor to maintain mental equilibrium in spite of many disturbing impressions and experiences.

– A picturesque description of this situation is to say that complexity *lies at the edge of order and chaos* (Waldrop 1992); a dialectician might call it a synthesis of order and chaos. A typical complex system encompassing order and randomness is also the terrestrial environment consisting of atmosphere and hydrosphere, which is governed by "orderly" laws but is subject to chaotic fluctuations going as far as hurricanes. Perhaps the best book on the interrelation between simplicity, chaos, and complexity is (Cohen and Stewart 1994).

– The element of *chaos* is by no means only negative: frequently it provides spontaneous novelty and creativity. Natural selection in Darwinian evolution is based on random mutations that occur spontaneously. (The "survival of the fittest" then restores order, possibly on a higher level.)

– The antithesis to Darwin's "struggle for survival" is "*cooperation for survival*", e.g. between algae and fungi to form lichens. Similar to cooperation is *adaptation* to the environment. An important example of self–organization, struggle, cooperation and adaptation is the *market economy* which is, however, beyond the scope of the present book. An exemplary case of cooperation between natural scientists and socio–economists is the Santa Fe Institute, cf. (Waldrop 1992). See also sec. 4.1.

– The concept of *emergence* of new unexpected features is essential. Examples are physical and chemical *phase transitions*, e.g. from ice to water and then to vapor; but also the emergence of life and the emergence of mind, cf. Fig. 2.13 on p. 56.

– There must also be a relation between complexity and the concepts of *information* or *entropy* (sec. 4.3). The more complex a system is, the more information is needed to describe it. There is also the concept of *computational complexity*. Details are still controversial, cf. (Zurek 1990), (Weizsäcker 1985, Chapter 5), and particularly (Gell–Mann 1994, Chapter 3).

Thus complexity is a "complex" collection of interesting ideas and mathematical models rather than a unified scientific theory such as, for instance, quantum mechanics. It is a field that contains many fascinating open problems.

4.3 Entropy, information, and evolution

Ein kleiner Ring
Begrenzt unser Leben
Und viele Geschlechter
Reihen sich dauernd
An ihres Daseins
Unendliche Kette.

Johann Wolfgang von Goethe

Entropy and the Second Law of Thermodynamics

For our purposes, we use the definition of *entropy* given by the great Austrian physicist Ludwig Boltzmann (1844–1906):

$$S = k \log W \tag{4.3}$$

where S denotes the entropy, k is a universal constant (the Boltzmann constant), log denotes the natural logarithm, and W is a quantitative measure of *disorder* of the system under consideration. A precise definition of W is rather technical, but we do not need it; the interested reader will find it in any textbook of physics.

Left to itself (without external influences), a physical system will always tend to increase its entropy:

$$\frac{dS}{dt} \geq 0 \quad . \tag{4.4}$$

The standard example is a scientist's desk on which papers and books are accumulating. With normal use, the books and papers become increasingly disordered (until the scientist gets disgusted and cleans the desk). The disorder W and hence the entropy S increase.

Another example: if we take a glass of hot water and a glass of cold water, and pour both into a larger glass, the temperature of the mixture very soon settles at a uniform lukewarm state. The mixture is less ordered than the originally separated amounts of cold and warm water whose temperatures were different, exhibiting more structure than the resulting uniform mixture. Again S has increased.

A similar case is the dissolution of a piece of sugar in a glass of water: the sugar solution possesses higher entropy (less structure) than the system consisting of the piece of sugar and the glass of water.

Thus, entropy always increases. The reverse processes never occur: the books and papers never order themselves, the lukewarm water never separates into a cold and a hot portion, and the sugar contained in the solution never collects to form again a piece of sugar surrounded by pure water (the case of crystallization of sugar in a saturated solution is another case).

The fact that the *entropy never decreases* spontaneously, as expressed by eq. (4.4), is called the *Second Law of Thermodynamics*. (The First Law is the conservation of energy.)

Negentropy and information

This process can only be reversed by applying some information. The desk becomes ordered again by a scientist knowing in which order the books and papers should be arranged.

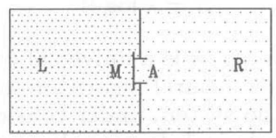


Figure 4.9: The shutter M is operated by Maxwell's demon

In the case of lukewarm water the situation is somewhat more complicated but very instructive (Fig. 4.9). As is known, molecules in a liquid or a gas perform a highly irregular and erratic movement. The molecules have individually different velocities, but the average velocity increases with increasing temperature. The higher the average velocity, the higher the temperature; in fact, the temperature is *defined* by the average velocity.

Let the container of Fig. 4.9 be separated by a wall with a very small aperture A. Let originally be the temperature the same (lukewarm) in both compartments L and R. The aperture can be shut or opened by moving the shutter M, operated by a small intelligent being, called *Maxwell's demon*. The demon observes the speed of the molecules approaching the aperture A. If a "fast" molecule approaches A from the right compartment R, he opens the shutter, and the molecule can pass to compartment L. If a "slow" molecule comes from R, the shutter remains closed. If a "fast" molecule approaches A from L, the shutter is also closed, but if a "slow" molecule comes from L, the demon opens the shutter. In this way, "fast" molecules will accumulate in the compartment L, and "slow" molecules will accumulate in R. Thus, the average velocity of the molecules in compartment L will become increasingly larger, and the average velocity will decrease in R. In other terms, water in compartment L will become warmer, and in compartment R, it will become cooler. It seems that the Second Law of Thermodynamics has been beaten!

The answer is that Maxwell's demon needs *information* about the molecules approaching the gate, especially about their speed, otherwise he cannot operate. In fact, the entropy difference between the initial uniform state S_0 and the final state S after the action of Maxwell's demon ($S < S_0$) can be considered as a measure of the total information I received by the demon in order to be able to operate:

$$I = -\Delta S \ , \quad \Delta S = S - S_0 \quad . \tag{4.5}$$

Thus, apart from a constant S_0, information equals a negative entropy:

$$\textit{information} = \textit{negentropy} \quad , \tag{4.6}$$

as expressed by L. Brillouin. An excellent biography of Maxwell's demon together with relevant original papers is (Leff and Rex 1990).

Genetic information

Life clearly also runs counter to the Second Law, both in an individual living being and in the course of biological evolution. Think of a newborn baby. It grows up, learns an enormous amount of things, and becomes a highly developed man or woman. In this course, the "order" or "organization" increases, entropy (disorder) decreases. Only after death, the Second Law takes over again: the body decomposes, disorder and hence entropy increase again.

Similarly, biological evolution has run counter to the Second Law: animals and plants become more and more complex and highly organized, from amoebae and bacteria to roses and human beings.

All this is far from being fully understood. Certain facts appear sure, however. First of all, the *fact of evolution* is not contested by any serious scientist. What is controversial is whether evolution proceeds in a completely random way, according to chance mutations and Darwin's "survival of the fittest", or whether it follows more or less an existing "cosmic blueprint" (Davies 1988). Probably the truth lies between these extremes. As Penrose (1989, p. 416) has put it:

To my way of thinking, there is still something mysterious about evolution, with its apparent "groping" towards some future purpose. Things at least *seem* to organize themselves somewhat better than they "ought" to, just on the basis of blind–chance evolution and natural selection.

The second fact is that the development of an individual is determined, at least principally, by giant molecules, called *genes*, which are contained practically unchanged in every cell of the individual's body. Genes are the basic carriers of genetic information which, so to speak, is *coded*, in very much the same way as literary information is coded in printed books or the performance of a Beethoven symphony is coded on a compact disc.

The complete genetic information is contained in a set of giant chain molecules, called *DNA* (nobody says *DesoxyriboNucleic Acid*); this set is also called a *genome*. They are Schrödinger's "aperiodic crystals" mentioned in the last section, and have the famous structure of a *double helix*.

We have the *Central Dogma of Molecular Biology* enunciated by Francis Crick (one of the co–discoverers of the double–helix structure of DNA) in 1970:

$$\text{DNA} \implies \text{RNA} \implies \text{proteins} \quad .$$

What is RNA? It is short for *RiboNucleic Acid*. RNA is closely related to DNA (DesoxyRNA): DNA, so to speak, contains less oxygen than RNA. DNA consists of long chains or relatively simple molecules called *nucleotides*, with their *bases*

$$\begin{array}{ll} \text{A:} & \text{adenine,} \\ \text{G:} & \text{guanine,} \\ \text{C:} & \text{cytosine,} \\ \text{T:} & \text{thymine.} \end{array} \tag{4.7}$$

Their exact composition is irrelevant for the present purpose, as well as the names: one always speaks of A, G, C, T only. RNA has the same constituents, but T replaced by

$$\text{U:} \quad \text{uracil.} \tag{4.8}$$

As an "aperiodic solid" (sec. 4.2), DNA is a very stable "original", or "template", from which RNA is obtained by "transcription", with the bases transcribed as follows:

$$\begin{array}{ccccc} \text{DNA:} & \text{A} & \text{G} & \text{C} & \text{T} \\ \Downarrow & \Downarrow & \Downarrow & \Downarrow & \Downarrow \\ \text{RNA:} & \text{U} & \text{C} & \text{G} & \text{A} \end{array} \tag{4.9}$$

The RNA thus obtained serves as "messenger" sent out by "her majesty" DNA in order to produce *proteins* which are well known to play a basic role as *enzymes* in the functioning of the living being. A good comparison is to regard DNA as the *legislative*, and RNA and the enzymes as the *executive* of the cell. A *gene* is that part of a DNA strand which is responsible for the production of just one protein:

$$\text{gene} \Longrightarrow \text{protein} \quad .$$

What is incredibly astonishing is that each gene, or each DNA, consists of the same four bases (4.7) only, whether DNA belongs to an amoeba, a bacterium, a rose, or any particular human person! (This, of course, is an overwhelming argument for evolution.)

As a matter of fact, the more complex the plant or animal becomes, the longer will in general be its DNA.

Proteins consist of *amino–acids*, which correspond to *triplets* (say UAG) of bases in RNA. A short but comprehensive and not too technical review is (Holzmüller 1984).

Problems and some preliminary answers

DNA is the basic and stable source of information for the formation of a plant or an animal. How this information leads to *morphogenesis* and hence to the formation of the fully developed organism, is a problem whose study has just begun. How does the same DNA give rise to such different cells as a cell of the liver, a blood cell, and a neuron? Another problem is the increase of information in DNA itself, leading to more and more complex organisms in the course of biological evolution.

There is an answer to such problems which, above all, has provided an attractive name and indicated future directions of research: *self–organization*. We have already treated it to some length in sec. 4.2, under the heading of "complexity", and we shall say more about it below, under the name "dynamic structures". (Repetitions from a slightly different perspective are intentional.)

How is the Second Law of Thermodynamics bypassed? Four partial but important answers have been suggested:

(1) *Metabolism*: the animal eats food (other animals, plants) which furnish not only energy but also (negative) *entropy*, i.e., information, since the food already has a rich structure. The body uses this information for its own organization and releases the food after having extracted energy and information, in a highly degraded (high–entropy) state into its environment (Schrödinger 1944).

(2) *Non–equilibrium thermodynamics*. The Second Law holds for states near thermal equilibrium. Animals etc. are in a state far fromthermal equilibrium, hence the Second Law does not immediately apply (Prigogine and Stengers 1984).

(3) *Enzymes* (organic catalyzers) may act as some kind of "Maxwell's demons" (Monod 1970; Eigen 1987).

(4) It seems that already in inorganic chemistry, *catalyzers*, enabling or accelerating a chemical reaction, may counteract the Second Law (Waldrop 1992, pp. 314–315).

Let us consider these questions and answers in some greater detail. (The rest of this section may be skipped.)

Nonlinear and dissipative structures

First of all, we must distinguish two types of structure: *autonomous structures* and *dynamical structures* (Mayer–Kuckuk 1989, p. 213). Autonomous structures, like crystals and DNA (Schrödinger's "aperiodic crystals"), are relatively permanent, their stability being assured by quantum mechanics; it is a *static* permanence not very much dependent on the external world. They are in a state of *equilibrium*.

Dynamic structures, on the other hand, change all the time, so their stability is *dynamic* rather than static. Think of yourself: you can only exist by the intake of food and liquids, which are used and, in a degenerated state, leave your system as waste. Dynamic structures, like convection cells in the inorganic world (Fig. 4.5 on p. 168), thus need interaction with the environment: heat in the convection case, and energy and negentropy in the organic case.

It is believed that these two cases, convection and life, have more in common than a superficial similarity. They are *dissipative structures*, *far from equilibrium*, which constantly absorb and dissipate energy (equilibrium is a state of minimum potential energy, so there would be no energy consumption there). Nonlinear dissipative systems frequently are *chaotic* (sec. 3.2). For instance, there is a close relation between heat convection (Fig. 4.5) and meteorological chaos of Lorenz type.

The thermodynamics of nonlinear systems far from thermal equilibrium was in particular investigated by I. Prigogine, who got the Nobel Prize in 1977. Synergetics (H. Haken) is closely related, and so is chaos theory (E. Lorenz), cf. sec. 4.2. This complex of models, theories and ideas seems to be responsible, at least partly, for the *dynamic structure* of living organisms.

The stable *autonomous structure* as represented by the genome (DNA) seems to serve as the *memory*, or library of permanent "legislative" information, as we have seen above. Another mechanism must be able to read and copy this information (DNA → RNA) and to convert the genetic "blueprint" into cell structures (RNA → proteins → structure). This "executive" is certainly a *dynamic structure*.

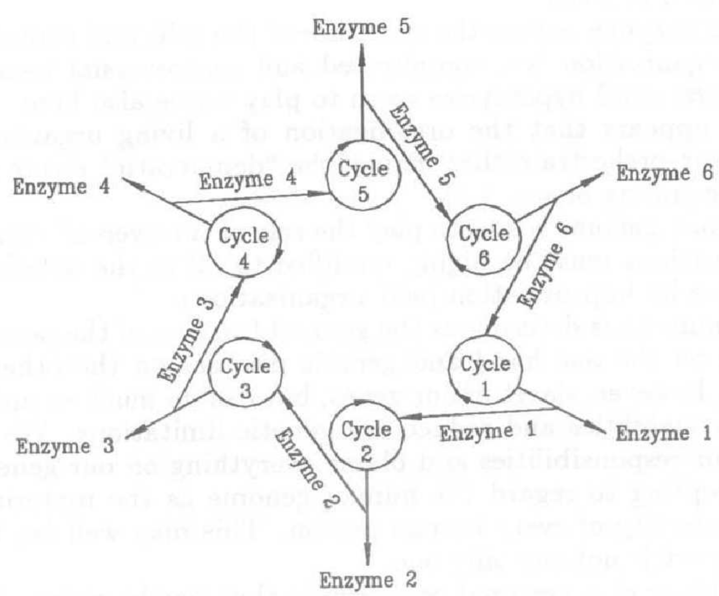


Figure 4.10: The hypercycle of Manfred Eigen (the number of cycles need not be 6!)

Nonlinear catalytic and autocatalytic processes generally play a great role. These processes are frequently "synergetic", that is, cooperative. An example is Manfred Eigen's "*hypercycle*" (Fig. 4.10).

A "cycle" may copy part of a nucleic acid (e.g. RNA) and produce a corresponding protein acting as an enzyme for increasing the activity of the *next cycle*. Thus the action of the hypercycle is one of mutual support of the work of the individual cycles. If Cycle 1 increases the production of its enzymes (Enzyme 1), this enzyme will stimulate the activity of Cycle 2, whose Enzyme 2 stimulates Cycle 3, etc.

Thus the cycles mutually help to increase their activity. It can be shown that also the stability is mutually reinforced. The cycles operate in an *autocatalytic* way, and also "assist each other" in a *cross–catalytic* manner. The hypercycle is a model for *cooperation* rather than competition in the organic world; cf. also (Jantsch 1980) and (Kauffman 1993). We also recognize a *feedback* acting within the hypercycle. As a matter of fact, the hypercycle takes energy and "raw materials" from the environment.

The great physicist John Archibald Wheeler was so fascinated by the hypercycle that he spoke of the "life machine" of Manfred Eigen (Wheeler 1994, p. 180)!

How the enzymes act on the structure of the cells and contribute to their self–organization is a complicated and controversial issue. Catalytic processes and hypercycles seem to play a role also here.

It thus appears that the organization of a living organism is of the conductor–orchestra rather than of the "democratic" string quartet type (see beginning of sec. 4.2).

The human genome seems to play the role of the over–all conductor, but the musicians must be highly qualified to fill in the details of the musical piece by improvisation (self–organization).

The genome thus determines the general features of the person, e.g. intelligence on the one hand and genetic diseases on the other hand. We are not, however, slaves of our genes, but can do much to make best use of our capabilities and reduce our genetic limitations. We cannot just shed our responsibilities and blame everything on our genes.

It is tempting to regard the human genome as the material basis of the self–identity of every human person. This may well be, but the material aspect is not the only one.

The stability of a personal organism is thus partly *static*. In part, stability in nonlinear systems may also be *dynamic*, that of a *fixed point* or *point attractor* (Fig. 4.11). (There are also more complicated "strange attractors" such as the Lorenz attractor well known from meteorology, cf. (Abraham and Shaw 1992), (Briggs 1992), (Lorenz 1993).)

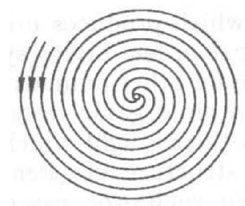


Figure 4.11: A point attractor. Every trajectory (path of development) ends at this "stable" point.

Metabolism and entropy

We now take up Schrödinger's ideas as elaborated by Prigogine. We are following the book (Schrödinger 1944) and the modern treatment of Schrödinger's ideas in (Götschl 1992, Part 2).

According to the Second Law of Thermodynamics, the entropy of a material system never decreases, cf. eq. (4.4):

$$dS \geq 0 \quad . \tag{4.10}$$

In a living being, however, the internal order increases, so that entropy decreases:

$$dS_i < 0 \quad . \tag{4.11}$$

There seems to be a contradiction.

In fact, we must consider the total system, organism(s) plus environment. The total entropy S_t must increase:

$$dS_t = dS_i + dS_e \geq 0 \quad . \tag{4.12}$$

A decrease according to (4.11) is possible if the entropy of the environment, S_e, increases even more strongly so that (4.12) is satisfied.

This is achieved by *metabolism*. Living organisms maintain their organization by "extracting order from the environment". (Quotations are from (Schrödinger 1944).) This is done by consuming food which already has a high degree of organization (plants, other animals).

The high order of an organism, as exemplified by its DNA, has an analogue already in the inorganic world: a crystal whose structure, as determined by quantum–mechanical laws, is already highly ordered. So a crystal, so to speak, is an inorganic forerunner of organic DNA, which Schrödinger recognized as an "aperiodic solid", of course of an incomparably richer structure.

This is a mechanism which produces *order from order*: higher degree order from lower–degree order. Living systems "feed upon negative entropy". Animals feed upon other animals or plants, but what about plants? They take order from organic rests contained in the soil (humus) but, above all, from the sunlight which furnishes energy but also has a rather low–entropy structure, cf. (Penrose 1989, pp. 319–321).

As Prigogine has shown, such processes occur far from equilibrium: think of a warm–blooded animal in winter. If this animal were in equilibrium with its environment, it would be frozen and hence dead. Energy, much energy is needed to maintain life, keeping the organism away from thermal equilibrium. (Consider only the high energy consumption of the modern world!)

Even inorganic systems can decrease their entropy (increase their order) through self–organization provided:

– they are far from thermal equilibrium;

– they are open, i.e., they can react with their environment;

– there is enough supply of energy and raw materials;

– there are auto–catalytic and cross–catalytic reactions such as Eigen's hypercycles (which are essentially nonlinear systems!).

Thus, just as DNA has its inorganic predecessor in crystals, metabolic processes have their inorganic forerunners in non–equilibrium thermodynamics and self–organization such as convection cells and the Belousov–Zhabotinsky reaction mentioned in sec. 4.2. Convection cells (Fig. 4.5 on p. 168) are a particularly clear example of order produced through the supply of thermal energy by heating.

It may even be that organic life is not an exception to the Second Law of Thermodynamics, but that the Second Law comprises two affirmations: an increase of entropy for ordinary "simple" systems and a decrease of entropy for "complex" systems including living organisms ((Waldrop 1992, Chapter 8), (Kauffman 1993, Chapter 8), for a related idea cf. (Weizsäcker 1985, Chapter 5)).

Anyway, Schrödinger (as most physicists in (Küppers 1987)) believes that new laws are at work in the organisms of biology. This does not mean a return to the old–fashioned "vital forces" but the new laws are "software laws" in the sense of sec. 4.5 (just as the income tax law serves as software law for the computer which calculated your income tax, operating, of course, according to the laws of physics).

Still, all these questions are extremely difficult, and the partial answers given so far are rather controversial. Fortunately, many of the pioneers have written excellent popular books: (Eigen 1987), (Eigen and Winkler 1975), (Haken 1981), (Kauffman 1993), (Prigogine and Stengers 1984), (Nicolis and Prigogine 1989), (Thom 1975), so that the reader can judge for himself. The extremes are marked by (Teilhard de Chardin 1955) and (Monod 1970): evolution according to a "holistic" plan, and completely random and mechanistic evolution. An excellent comparative treatment of all these various tendences and directions of research is found in (Moser 1989, Chapter 4). Balanced and concise presentations are given in (Davies 1988) and (Mayer–Kuckuk 1989, Chapter 9). The little pioneering book (Schrödinger 1944) has retained its charm and freshness for over half a decade and is my favorite. Off the beaten track are the passages on the genetic code in (Hofstadter 1979) and in (Cohen and Stewart 1994), other favorites of mine. The presentation of Sheldrake (1981) is unconventional but very readable and interesting. An excellent general introduction into evolution is (Edey and Johanson 1989), combining readability with a high level. If you look for open and controversial problems, consult (Duncan and Weston–Smith 1977). A monograph representing the current state of research on complexity is (Zurek 1990), a publication of the Santa Fe Institute, whose work is excitingly described by Waldrop (1992; general introduction) and Lewin (1992; emphasis on biology). Cohen and Stewart (1994) show how complex evolution really is.

4.4 Data and errors

Errare humanum est.

Latin proverb

The only quantity that is exactly measurable is a *number of individuals*, which is a positive integer (or zero): 0, 1, 2, 3, ... In fact, if a basket contains 5 apples, it will be exactly 5 and not 4.9937.

All measurements of quantities which are permitted to assume a continuous range of values and which are expressed by real numbers, *are only approximate*. Even the simplest geometric quantity, the distance between two points in "our" Euclidean three–dimensional space, can be measured only approximately, as we have discussed at length in sec. 2.4.

The reason is that precise definitions and exact laws exist only in mathematical logic and in mathematics (as far as Gödel permits!). In nature, concepts are always inexact. In physics, the way out is to make the experimental arrangement as exact as possible, and then, by a leap of faith, *idealize* the situation and postulate that the measured values obtained in this way are absolutely accurate. This may or may not work.

One of the most accurately known physical constants is the speed of light c in vacuum. Its value is

$$c = 299\,792\,458 \text{ m s}^{-1} \quad , \tag{4.13}$$

that is, meters per second. The measuring accuracy is so high that the last three numbers might be 459, but not, for instance, 464. We say that c has been measured to an accuracy of ± 1 m s^{-1}. This is the standard error or r.m.s. (root mean square) error σ which is defined statistically and known to everyone who has ever handled empirical data. Now, $\sigma = \pm 1$ m s^{-1} does not mean that only 457 or 459 may be possible outcomes of measurement; even an error of 3σ (possible outcomes 455 or 461) may occur, but with rather small probability. (Even larger deviations are theoretically possible, but extremely improbable.)

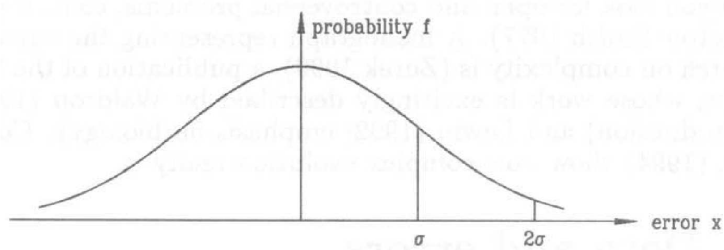


Figure 4.12: The Gaussian error curve

This is a consequence of the famous Gaussian error curve shown in Fig. 4.12. The equation of this curve is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{x^2}{2\sigma^2}} \tag{4.14}$$

where π and e are standard mathematical constants. (Actually f is not a probability but a probability density, but this distinction, important as it is, is irrelevant to the present discussion.)

This symmetric figure clearly shows that positive and negative errors are equally possible, and that "outliers", say $x = 5\sigma$, are theoretically possible although their probability is so small that they are considered practically impossible.

Important is the concept of *relative standard error*

$$\frac{\sigma}{c} = \frac{1 \text{ m s}^{-1}}{299\,792\,458 \text{ m s}^{-1}} \doteq 3 \times 10^{-9} \text{ (dimensionless!)} \quad . \tag{4.15}$$

This accuracy is considered almost the highest accuracy that can conceivably be achieved. So why not consider it "absolutely accurate"?

The chief limitation of the accuracy of

$$c = \frac{\text{length}}{\text{time}} \tag{4.16}$$

is the limited accuracy of measuring length or distance, that is, the *definition of the meter*. The limited accuracy is due to the difficulties mentioned at the beginning of sec. 2.4 regarding point definition and distance measurement. The meter was defined as a certain multiple of the wavelength of monochromatic light of a certain frequency, and it was simply not possible to increase the accuracy of the frequency definition. On the other hand, time definition by means of an atomic clock is considerably more accurate.

So, some twenty years ago, it was in fact decided to give up the old definition of the meter and define length as time $\times$ c. Thus the *defining constants* were taken the second and the velocity of light, with c according to (4.13) *errorless* and fixed once and for all. The meter is now a *derived* quantity: it is the distance covered by light in c^{-1} seconds,

$$1\,\text{m} \,\hat{=}\, \frac{1}{c}\,\text{s} \quad . \tag{4.17}$$

Thus, in certain rare cases, it is permitted to consider an accurately measured value as absolutely errorless, but then it may be necessary to redefine other quantities in such a way that logical (and numerical!) contradictions are avoided.

It would not be permissible to regard all three measured angles of a triangle as errorless, since this results in a numerical contradiction: the sum of the three angles in a triangle must be 180° or π:

$$\alpha + \beta + \gamma = \pi \quad , \tag{4.18}$$

and this condition is not in general satisfied by the measurements. Of course, one could avoid this difficulty and measure only the angles α and β, computing γ from (4.18). Similar procedures are applied frequently enough, but may not be quite fair for two reasons:

(1) Measuring α, β and γ provides a useful check to avoid gross errors, e.g. $\alpha = 74°13'35''$ instead of the value (2.14).

(2) All three angles should have "equal rights" to be measured, which is a principle of symmetry (or democracy if you prefer).

Thus, honestly, we should measure all observable quantities and remove any numerical contradictions by a *least–squares adjustment* as mentioned in sec. 2.4 on p. 41.

This "honesty" is rewarded: the study of measuring errors, due to R. Bošković (1711–1787), A.M. Legendre (1752–1833) and, above all, to C.F. Gauss (1777–1855), the "princeps mathematicorum", helped found mathematical statistics. Least–squares adjustment was also elaborated by Gauss and led to the beautiful geometrical interpretation in terms of normal projection in a higher–dimensional space mentioned in sec. 2.6.

4.5 Complexity and reductionism

If physicalism is correct, then even a family of monkeys in a tropical forest is "in principle" a solution of the Schrödinger equation.

Carl Friedrich von Weizsäcker

We are resuming the study of complexity, started in sec. 4.2, from a different perspective.

In biology there have been essentially two opposite opinions:

(A) *Vitalism*: A living organism is not determined by the laws of physics only; there exist special "vital forces" which cause the purposeful behavior of living organisms, their special structure, their ability to heal wounds and even to regenerate lost organs (this is in particular conspicuous in lower animals such as polyps or starfish), etc.

(B) *Reductionism*: a living organism is nothing else than a very complex and well–structured system, which is completely governed by the ordinary laws of physics and chemistry. Since chemistry, through the laws of quantum mechanics, is thought to be reducible to physics, also the laws governing the apparently so special behavior of living organisms are reducible to the laws of physics. This is reductionism or physicalism.

The main empirical data are clear:

(1) The behavior of animals and plants is completely different from any mechanisms or similar man–made automata.

(2) All physical experiments performed with living organisms or with living tissue have never indicated any measurable deviations from the ordinarily known laws of physics and chemistry.

(3) There seems to be no sharply defined boundary between highly–organized macromolecules and the most elementary organisms, between chemistry and biology.

It is safest and least controversial to consider living organisms as very elaborate and highly organized complex systems. (Is a beautiful mountain flower *only* a complex physical system? Between ourselves, my answer is "no", but don't tell this to anybody!)

Complex systems

It is instructive first to consider a highly complex modern electronic computer. Is it governed by the laws of physics? Undoubtedly, *yes*.

Is it *fully* governed by the laws of physics only? *No*. Let me try to explain.

Even in classical mechanics we need not only Newton's laws, expressed by ordinary differential equations of the second order, but we also need *initial conditions* (position and velocity at a time $t = t_0$), to get a well–determined motion. Boundary conditions are essentially the same as initial conditions.

For a computer, the initial and boundary conditions are nothing else than its input: program and data, in other terms, the *software*. The computer with zero input constitutes the *hardware* (thus, for the present purpose, hardware also comprises fixed programs built into the computer).

With no input, the computer will not work in the sense of producing a useful output. As soon as the software is introduced, the computer starts to work and produces a useful output.

The input may consist of a complicated program *which in itself may contain a law*: for instance a sequence of highly complex mathematics. Or the program for computing income tax may contain a mathematical–logical form of the income tax law (Davies 1988, p. 144). Hence we may well speak of *software laws*.

Thus, the operation of a computer is governed by the physical hardware laws and the non–physical software laws! (Nobody would claim that the income tax law is derivable from the laws of physics . . .)

Thus the laws governing the behavior of living organisms may well be regarded as some kind of *software laws describing biological complexity*, following again (Davies 1988, p. 142). (This book is my favorite reference for the present problems.)

Thus the "biological laws" define the *initial or boundary conditions* for the work of our "biological" systems.

This comparison between animal and computer, imperfect as it is, shows clearly not only the possibility, but even the necessity of laws other than physics to describe the activity of a computer and, *a forteriori*, of a living organism. Note that in sec. 4.1 we have used the computer in a similar way, as a model for mind–brain interaction.

The old fight between vitalists and reductionists is thus raised to an objective and unemotional level. Any molecular biologist who speaks of "vital forces" would commit scientific suicide, but if he (she) speaks of "software laws", nobody will pay attention to the fact that, possibly, the same thing is expressed in two different forms.

Software contains *information*. Thus information must play a decisive role for biological systems. Perhaps oversimplifying, we may thus say

$$\text{life} = \text{matter} + \text{information} \tag{4.19}$$

(Küppers 1987, p. 17). This is probably true, but what is information? Unfortunately it is by no means a clearly and unambiguously defined concept. The relation between entropy and information outlined in sec. 4.3 is rather generally accepted, but neither gives a complete definition nor is understood by all scientists in the same way, cf. (Weizsäcker 1985, Chapter 5). It is one of the basic paradoxes of science and philosophy that the most fundamental concepts such as matter, mind, or information are so ill–defined.

By regarding a living organism as a *complex system* governed by the laws of physics and "software laws" containing information we have given the problem a simple (though probably oversimplified) structure which may serve as a basis for more detailed investigations. The related concept of *self–organization* (sec. 4.2) may serve for similar purposes.

Using the terminology of Niels Bohr, physical and biological laws (if you don't like this term, speak of software laws) are *complementary* (very much in the sense of complementary subspaces which are orthogonal to each other, cf. Fig. 2.19 on p. 66 and Fig. 6.1 on p. 246).

This has been drastically formulated as follows ("*Bohr's paradoxon*"): in order to determine whether a cat is fully governed by physical laws, it is not sufficient to determine its weight or its bodily temperature. One must use an X–ray equipment, which has to be very powerful to determine the cat's exact internal structure, so powerful that it may well kill the cat or damage it irreversibly. This is not sufficient, however: to get other physical parameters, we must implant physical equipment in the cat's body, and finally we must dissect it. By then, the cat is surely dead. Thus, life and a "full" physical examination are incompatible with each other!

A living organism is an individual "whole". This is what the concept "*holism*" means: the whole is more than the sum of its parts. Since we have already used this concept before (secs. 2.1 and 3.5), it may be appropriate now to remind the reader again that "holism" is not a mutilated version of "wholism" but comes from the Greek word "*holos*" which means "entire" or "whole" (after all!).

Thus "holism" and "reductionism" are complementary terms; this is beautifully expressed in the drawing in (Hofstadter 1979, p. 310 and the subsequent "ant fugue").

On reductionism

As we have said at the beginning of this section, it seems that the laws of physics hold also for living organisms. Let us consider this question more closely. Physical laws have an inherent inaccuracy as we shall see in sec. 6.5. Experiments so far have shown that they are satisfied in plants and animals *within the measuring accuracy*; this assertion we shall call *moderate reductionism*. Furthermore, as we have just seen, more precise physical measurements may interfere with life (of a cat, say).

So the question whether physical laws are "really" and "absolutely" the same in living and nonliving matter may well be meaningless. Nevertheless, this has been frequently asserted, so we shall call it *strong reductionism*. If someone asserts this, we can immediately retort: reduction to which physical laws? The "real" ones if they exist? The laws found in our books of physics? But does Nature, or God, read those books? Questions are becoming cynical, so let us stop them.

Strong reductionism is asserted mainly by biologists. Biologists often have a much stronger faith in physics than physicists themselves. (Similarly, physicists have a much stronger faith in mathematics than mathematicians, cf. the end of sec. 2.3. This seems to be due to inside knowledge: it is said that Napoleon's valet had much less respect of him than most other people.) The book (Küppers 1987) contains highly interesting articles by great contemporary physicists on this topic which are very relevant to our discussion. Great physicists are usually rather cautious regarding reductionism. An eminent physicist such as Walter Elsasser speaks of "biotonic laws".

The old opinion of Whitehead (1925, p. 115–116) is still of interest in this respect:

> The concrete enduring entities are organisms, so that the plan of the *whole* influences the very character of the various subordinate organisms which enter into it. In the case of an animal, the mental states enter into the plan of the total organism and thus modify the plans of the successive subordinate organisms until the ultimate smallest organisms, such as electrons, are reached. Thus an electron within a living body is different from an electron outside it, by reason of the plan of the body. The electron blindly runs either within or without the body; but it runs within the body in accordance with its character within the body; that is to say, in accordance with the general plan of the body, and this plan includes the mental state.

This is holism expressed in beautiful English.

Whitehead's position does not need any "vital forces" but may rather be regarded as imposing either constraints in the sense of sec. 3.1 (cf. eq. 3.13 on p. 75) or boundary conditions in the form of "software laws". To fully demystify "biotonic laws", compare them to the income tax law as mentioned above ...

But let us assume now that even strong reductionism is true. Thus, all physical laws are perfectly valid even in a living organism. This includes the laws of quantum mechanics, and the "orthodox" Copenhagen interpretation (sec. 3.5) involves the mind of the observer. Thus, via quantum mechanics, life, in being reduced to physics, brings mind in again, especially according to the view of Eugene Wigner (p. 105).

This is probably the last thing which reductionists want, but it gives a nice dialectic loop (Fig. 4.13) This shows where logical reasoning may lead. I do not necessarily agree, but this may be a profound truth in the sense of Bohr, of which the opposite is also a profound truth (sec. 2.5). Here again, we should beware of oversimplification, of Whitehead's "fallacy of misplaced concreteness"!

Reductionism versus constructionism. Reductionism says that if we investigate living beings as regards their physical properties, we shall find that the usual laws of physics, as we know them, apply to this process. Quite another problem is whether we can construct an animal, say, from the laws of physics only. Thus, *constructionism*, so to speak, is the *inverse problem* (sec. 3.8) of reductionism.

To clarify this, let us consider the following example. A small child takes apart an old–fashioned mechanical watch and announces proudly

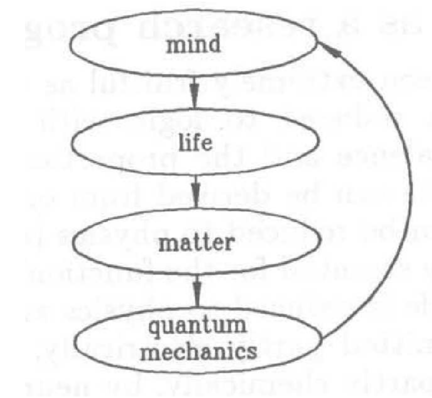


Figure 4.13: The self–reference loop of reductionism according to E. Wigner

that the watch is nothing but cogwheels, springs, screws, pinions, and other simple mechanical parts. The father is upset and tells the child to put these "simple" mechanical parts together again to obtain the original watch. The child is unable to do this. "Taking apart" was a reductionist procedure, and "putting together" is a constructionist problem which is much more difficult: the inverse problem is usually essentially more difficult than the direct procedure.

In 1972 Philip Anderson asserted that "the reductionist procedure does not by any means imply a 'constructionist' one: The ability to reduce everything to simple fundamental laws does not imply the ability to start from these laws and reconstruct the universe" (quoted after S.S. Schweber, Physics Today, November 1993, p. 36).

H. Primas formulates it even more strongly: "Every machine relies for its operation on the laws of physics and chemistry, but the machine's design is a higher–order principle. In this sense machines are irreducible to physics" (quoted from Moser 1989, p. 137).

A rather generally prevailing compromise might be formulated as follows. Physical laws hold throughout nature, living and lifeless alike. What is characteristic for biology are not the *laws* but the *boundary conditions*, which are responsible for the complexity of living organisms, the physical laws being quite simple. You may call these boundary conditions "biotonic" laws or "software laws", depending on your preferred way of thinking. Even the freedom of the will may be related to boundary (or initial) conditions, cf. sec. 6.4 and Fig. 6.1.

Reductionism as a research program

Reductionism has been extremely fruitful as a method of research. Mathematics has been reduced to logic with considerable success (sec. 2.1). Chemical valence and the properties of the periodic system of chemical elements can be derived from quantum mechanics; in this sense, chemistry can be reduced to physics (sec. 3.5). Physics and chemistry are absolutely essential for the functioning of a plant or of an animal. In this sense, life is reduced to physics and chemistry. Signals in the brain are transmitted partly electrically, or electrochemically, along the neuron, and partly chemically, by neurotransmitters, across the synaptic cleft (sec. 1.1). In this way, thinking may be said to be reduced, at least in part, to physics and chemistry. Last but not least, the study of atoms and molecules is reduced to the theory of elementary particles and their gauge theories: the Standard Model and, perhaps, superstring theory as a Theory of Everything (secs. 3.6 and 6.6).

It can safely be said that most research activity in the natural sciences is concerned with reductionist work. Thus reductionism has been immensely successful as a research program; cf. (Popper 1982) and the recent books (Cohen and Stewart 1994) and (Gell–Mann 1994).

Is reductionism the final philosophical answer to all problems of natural science? What we have just tried to show is that the answer is a clear "No"; Popper (1982, pp. 131–132) confirms this. *A reductionist analysis may lose basic information*, cf. "equation" (4.19).

Equally important conceptually are complexity theory, holism, and constructionism. Thus we have three antitheses of complementary concepts:

simplicity and complexity: (Cohen and Stewart 1994),
(Gell–Mann 1994),

holism and reductionism: (Hofstadter 1979, pp. 310–336),

reductionism and constructionism: discussed above.

A research program much smaller than reductionism but of increasing importance is *complexity theory*, cf. sec. 4.2. Reductionism goes "down" from the complex to the simple, whereas complexity theory goes "up", finding perhaps a new higher simplicity emerging from complexity, "order out of chaos", thermodynamics emerging from statistical mechanics, the DNA of life emerging from organic chemistry by Eigen's "life machine" (Fig. 4.10 on p. 179).

Holism is less well regarded by science, which is analytical rather than synthetic, and vitalism is completely out of fashion at the present time. (The famous biochemist and Nobel laureate A. Szent–Györgyi once wrote: "When a molecular biologist calls you a vitalist it is worse than when an FBI man calls you a Communist" (quoted after W.M. Elsasser's autobiography).)

Quine's model

Some positivists may tend to consider logic as "true in all possible worlds" (Leibniz), mathematics is reduced to logic, physics is expressed in mathematical laws and hence reduced to mathematics, chemistry is reduced to physics, and biology is reduced to physics and chemistry, and hence to physics. This throughgoing reductionism" has been seen to be endangered by logical paradoxes, Gödel's theorem, Heisenberg's uncertainty relations, and by inaccuracies of both measurements and theories in general, not to forget self–reference (Fig. 4.13).

Quine (1961, p. 42) has given a somewhat different and probably more realistic picture (Fig. 4.14).

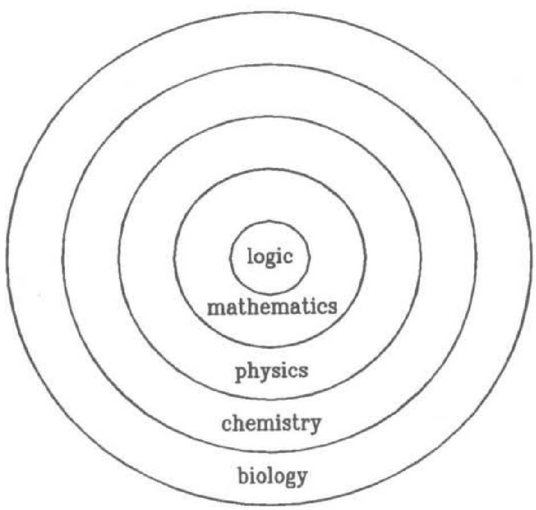


Figure 4.14: Quine's model

Empirical observations may show that existing theories may require modification. Then it is advisable to start changes at the periphery, in biology. If necessary, then chemical laws may be changed, then physical laws. If still necessary, however, even mathematics and, as a last resort, logic are not immune against change.

Thus, according to Quine, even logic and mathematics are not given *a priori* once and for all, as gifts from heaven, perfect and changeless.

In fact, Weizsäcker (1985, p. 313–319) has worked on a "quantum logic" in order to improve our understanding of quantum theory (cf. end of sec. 2.6), and also Weyl's opinion (the sentence following the quotation at the end of sec. 2.3) fits well into this scheme:

A truly realistic mathematics should be conceived, in line with physics, as a branch of the theoretical construction of the one real world, and we should adopt the same sober and cautious attitude toward hypothetic extensions of its foundations as is exhibited by physics.

Part C

Philosophy

Chapter 5

Philosophy for scientists

5.1 Realism, idealism and dualism

Words, words, mere words, ...
William Shakespeare, Troilus and Cressida

Now we come to the difficult and treacherous area of philosophical "isms". These terms are by no means clearly and uniquely defined. They are frequently loaded with emotions and used, not only for objective classification, but also to fortify one's own position and to depreciate the other's opinion. Especially the words "materialism" and "idealism" are used for such purposes. The argument partly repeats sec. 1.2 in a somewhat broader context.

Realism and idealism

Naive realism. We see the world as it is. If we see a house over there, we can be sure that it is there. If we see a car approaching us at high speed, we better give way in order not to be hit. This naive realism is so obvious, so appropriate to our daily life, that it appears unnatural to doubt it, and it also leads to science: physics, chemistry, and biology.

Now, however, we get a "*dialectic reversal*" (sec. 2.5). Science tells us that we do not see the house as such, but the walls of the house reflect electromagnetic waves (light coming from the Sun), and these waves, after having hit our eye, produce an image on our retina which is processed by our brain (sec. 1.3). So what comes from the external world, is only an image on our retina, which is a mental phenomenon.

Hence what is primary is our *mind*, and we do not see the house, we see some mental image, which could also be produced in another way, by a motion picture or by a television program. Thus this leads to

Idealism. The primary perceptions are affections of the mind, sense data, from which by thinking we reconstruct an external world which may not even exist in the way it appears.

Now, however, we ask: where do these sense data come from? Unless we suffer from an illusion, sense data must come from the external world (of which also cinemas and television sets are part).

After naive realism as thesis and idealism as antithesis, we thus arrive at a synthesis:

Critical realism, or *scientific realism*. Through our senses we get information on the external world, which is certainly partial and imperfect (we do not see ultraviolet and infrared light, for instance) but is essentially true. If this information is not true (illusions, cinema) we at least understand why it is not true: science (physics, evolutionary theory of knowledge, psychology, and psychiatry) tells us about this.

This is the scientific world picture, which essentially gives the same results as naive realism, but in a refined way. It is maintained by practically all scientists and philosophers, even by those who do not admit it.

Two remarks are in order, however. This dialectic game could be continued, getting a refined idealism, an even more sophisticated realism, etc. (This is beautifully shown in Fichte's Wissenschaftslehre of 1904, cf. sec. 5.3.) We could stop at any stage of realism or idealism, and many philosophers did stop at an idealistic position. We shall be satisfied, however, with scientific realism, at least as a working hypothesis.

The second remark is the following. Neither realism nor idealism, at any stage, can be proved or refuted. If we take an extreme idealistic position, we arrive at

Solipsism. Only my own ideas are real, my sense perceptions are nothing but illusions, I live exactly like in a dream. ("*Solus ipse*" means: only I exist, all the rest is illusion.)

It is impossible to refute this position. (When I heard this for the first time at the age of 16 years from our excellent philosophy teacher, I wanted to test it. Waking up one morning, I decided that "in reality" I continued to sleep and only dreamed that I had woken up. I persisted in this attitude, behaving normally but considering everything a dream: going to school, studying, taking examinations, playing etc. It worked perfectly and was absolutely self–consistent. After a few weeks, however, I decided that I had done my solipsist homework and that I had grown tired of it, and continued to live normally.)

As I said, it is difficult to refute the solipsist position, but it is still more difficult to maintain it. Bertrand Russell gave a nice example: at a philosophical congress, a solipsist philosopher criticized a colleague, asking him why he was not also a solipsist. (This, of course, is self–contradictory, if not in theory, then certainly in practice.)

Another self–refuting attitude is universal

Skepticism. The skepticist doubts of everything. If his skepticism is true, then the proposition "My skepticism is true" cannot be subject to doubt, making impossible the skepticist's intention to doubt *everything*. This "dialectic reversal" is of course dangerously close to the paradox of the liar. For a comprehensive discussion of this and related matters cf. (Stegmüller 1969).

Materialism and science

For those who want to eliminate the last trace of thinking or mind from the world, realism becomes materialism. Only matter is real, mental activity is only an illusion. (Here I do not include dialectic materialism, which will be discussed below.)

The problem is only to define what matter is. An obvious answer is obtained by kicking a stone: a solid object.

Modern physics gives a more sophisticated answer. As we have seen in sec. 3.6, a hydrogen atom consists of a nucleus (proton) of about 10^{-15} m, surrounded by the orbit of an electron which has a distance of about 10^{-10} m. To understand this proportion, assume that the nucleus has the size of the Earth. Then the distance to the electron is greater than the distance from the Earth to the Sun! So matter consists almost exclusively of empty space . . . If you kick a stone, you are kicking empty space or rather, an electromagnetic field: only the strong atomic and molecular forces provide the appearance of solidity.

It was shown only by modern physics how difficult a definition of the concept of "matter" is. It is furthermore – explicitly or implicitly – dependent on one's philosophical views: "Tell me what is your philosophy, and I will tell you how you must define matter" (Weizsäcker 1985, p. 165). The only generally accepted definition is very abstract: "Matter is what satisfies the laws of physics" (Weizsäcker 1977, p. 586).

Let us illustrate this by classical physics. Newton's law of motion

$$m\frac{d^2\mathbf{x}}{dt^2} = \mathbf{F} \quad , \tag{5.1}$$

where m denotes the mass (the "amount of matter"), $\mathbf{F}$ represents the force, and $d^2\mathbf{x}/dt^2$ (second derivative of the position vector $\mathbf{x}$ with respect to time) is the acceleration.

Now we could argue: this is a mathematical equation connecting mathematical symbols, and mathematics belongs to some "ideal world". So physics must be interpreted in the sense of idealism. This argument has been advanced, especially with the much more difficult and abstract equations of relativity and quantum theory.

It is easy to give a counterexample: 3 apples + 2 apples = 5 apples, and it does not follow that apples are "unreal" mathematical concepts. Mathematics does apply to the real world, so it does not permit us to use it for deciding between realism and idealism.

Nevertheless we have seen in secs. 3.4 and 3.5 that matter is not at all a simple concept. It is related to space–time curvature or satisfies Schrödinger's equation in "infinitely–dimensional" Hilbert space. Thus it has become very abstract. If we adhere to a scientific realism (as the majority of physicists do, including the present author), the standard "Copenhagen interpretation" somehow introduces the mind of the observer, and hence an element of idealism. (This is not so terribly surprising if we remember that, at the beginning of the present section, we have obtained "scientific realism" as a higher–level synthesis of "naive realism" and idealism.)

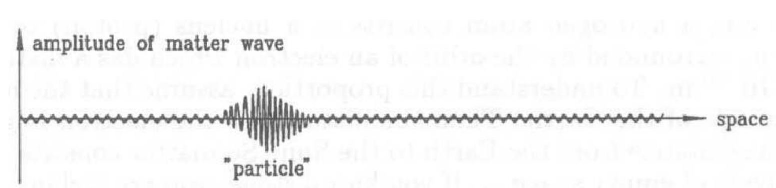


Figure 5.1: A particle as a wave packet

Anyway, matter has become very abstract indeed. Quantum mechanics regards a "particle" as some kind of "wave packet" (Fig. 5.1). Such a wave packet is not an exact point (where the "particle" is supposed to be), it has a certain finite though very small size. The important thing, however, is that the wave function is concentrated at the particle, but it has a non–vanishing, though very small, amplitude *at all* *points of space*. So to speak, a particle is everywhere! The Newtonian belief in strictly local material points, or even strictly localized extended material objects, is called by Whitehead (1925, p. 72) the "*fallacy of simple location*". Whitehead arrived at this conclusion by other arguments, some philosophical. Another physical argument against simple location comes from general relativity: mass points are singularities of curvature of space–time, but change the curvature also at other points.

In sec. 3.4 we have seen that eq. (3.55) on p. 96 provides a logically particularly beautiful, though practically hardly useful, definition of matter in terms of geometry.

The mathematical structure of physics is the same, whether we adhere to materialism or idealism. Thus, from a higher logical point of view, the distinction between materialism and idealism may mainly be one of words or of language, highly colored by emotions (if I hurt my foot by kicking a stone, I may not be very satisfied by the consideration that anyway, only "one mathematical equation has kicked another mathematical equation" ...). Bertrand Russell calls the philosophical basis of physics a "neutral monism", being neither materialism nor idealism (or both).

Monism and dualism

Besides realism (or materialism) and idealism, we have another important pair of more or less opposed "isms": monism and dualism.

Dualism, in its clearest and sharpest form, goes back to René Descartes (1596–1650). He considers *matter* and *mind* as two totally different and almost completely unrelated "primitive concepts" (philosophers give them the lofty name of *substances*). So to speak, this is the Christian or rather Platonic idea of "body" and "soul" in its extreme philosophical abstraction. By radically separating matter and mind, Descartes cut himself off from any relation between these two "substances". In this way he artificially created a problem whose solution has kept busy generations of philosophers. It is not surprising if an artificial problem invented by a genius has given rise to even more artificial, though ingenious solutions.

The greatest problem was the essential difficulty of interaction between the two "substances" ("noninteraction" or "independence" was considered a main property of a philosophically respectable "substance"). One of the most curious solutions was the following. Compare matter and mind to two parallelly running clocks. The easiest way to keep them synchronized is to connect them in some suitable manner. This solution was excluded by the supposed impossibility of interaction of the two "substances". Thus there are two other possibilities: either God has created both clocks so perfectly that they always remain absolutely synchronized, or God is constantly watching them to keep them continuously synchronized, immediately correcting any deviation. This is a fine example of where "sharp" or "exact" philosophical reasoning can lead to. (God as the "Universal Metaphysical Problem Solver", cf. sec. 1.2.)

The old philosophical fallacy was the belief that one has to start from "clear", "true", "evident" and "precise" principles ("premises") and then follow sharp logical reasoning *wherever it would lead you.* Only a slight error in the premises might lead to large errors in the result ("ill–posed problem", see below). Whitehead (1925, p. 75) calls this unjustified sharpening of concepts ("noninteracting substances"!) the "*fallacy of misplaced concreteness*". Ivan Supek (quoted in the introduction) is more direct:

> The old metaphysicians got caught in the trap of some absolutized words and concepts, categories or principles; and wishing to construct a consistent [philosophical] system, they locked themselves into a lifeless ivory tower.

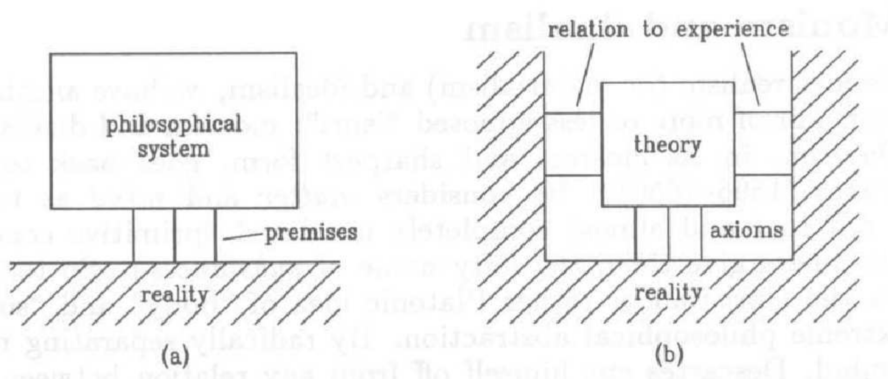


Figure 5.2: Classical philosophy and the axiomatic method

Relation to the axiomatic method. This is formally similar to but in reality quite different from the modern axiomatic method (sec. 2.2). Here one also starts from premises (axioms) which, however, are not necessarily considered "true" and "evident" but assumed tentatively. Then, again, a logical edifice (theory) is built by rigorous derivation from the axioms. *Then*, however, it is tested whether the theory fits experience (horizontal connections in Fig. 5.2(b). If it does not fit, other axioms are tried, etc.

Thus the basic difference between these two approaches is

(a) Premises are considered "absolutely true" and all logical consequences (however absurd) are accepted.

(b) Premises (axioms) are accepted preliminarily and the logical consequences are tested by comparison with reality.

It is clear that method (b) can be much better fitted to reality than method (a), which, already from the picture (Fig. 5.2(a)), appears much less "stable".

In the language of sec. 3.8, approach (a) may be compared to an ill–posed (unstable) problem, whereas (b) may be regarded as a well–posed and hence stable problem.

After this digression let us now return to dualism and monism.

Modern dualism or interactionism. This point is well argued and expressed in (Popper and Eccles 1977). There is no reason why mind and matter (our body) should not directly interact. A bad tooth (material) may cause pain (mental). This causal connection is almost trivial. But mind can also act on matter: a toothache can cause me very quickly to see a dentist, which otherwise I should have done only half a year later. Considering mind as a higher level than matter, this is an example of *downward causation* (sec. 4.1).

Monism. The doctrine that there is only one substance goes back to Parmenides (5th century B.C.), was represented by the Neoplatonist Plotinus (204–270) and revived by Baruch Spinoza (1632–1677). Albert Einstein is known to have been a great admirer of Spinoza.

Spinoza taught that there is only one substance (Greek: *monos* = one): *deus sive natura*, God or (equivalently) nature. This doctrine is also called *pantheism* (all is God).

This one substance has two "sides" or "attributes": matter and mind (and many others since God must be infinite in every respect). This view solves the problem of interaction of mind and matter, because, substantially, both are the same.

Spinoza's view, as those of Parmenides and Plotinus, was typically *holistic*: all essentially is one.

Russell's neutral monism. The general attitude of Russell is *logical atomism*, the very opposite of holism. Nevertheless to him, the (essentially atomistic in the sense of discrete) sense data are primarily neither mental nor material: They are "monadic structures" (see sec. 5.3) and, depending on the arrangement one gets subjects or "minds" (formed of all data collected by a person) and "material" objects (formed by all data that are coming from a given object). This definition, seemingly circular, can be made logically acceptable, breaking the "vicious circle", in much the same way as in Russell's definition of natural numbers (positive integers, cf. sec. 2.1).

Materialism and idealism both belong to the category of monism. Therefore both Hegel (a dialectical idealist) and the dialectic materialists thought very highly of Spinoza's monism, which could easily be considered a forerunner of materialism (in *deus sive natura*, consider "God" just an example of extravagant terminology, retain nature and call it matter!). On the other hand, explicit idealism and especially dualism are much less acceptable to dialectic materialism, because the supposedly Christian notion of "immortal soul" may be hidden below.

Panpsychism is another modern form of monism in the sense of Spinoza. Mind is simply the inside, and matter the outside of everything. This doctrine has been more or less explicitly favored by so different philosophers as Whitehead, Teilhard de Chardin, and also in a sense by Russell as we have seen. It appears very plausible if regarded from the point of view of human perception and thinking: mind "looks" from the inside at the outside material world. It seems to run into difficulties at the atomic level: does every atom (in a stone, for instance) also have an "inside" of a mental character? Leibniz (1646–1716) affirms just this and calls these "mental atoms" *monads* (cf. sec. 5.3). A similar but more modern version of such a theory was held by the early Russell and greatly elaborated by Whitehead. To be sure, it looks rather extravagant, but is fascinating and may contain an important element of truth (we already know that Whitehead once said: "It is more important that a proposition be interesting than that it be true"). The possibly false theory of monads is indeed more interesting than the certainly true relation "1+2=3" ...

Panlogism (Greek: all is logic) is a label frequently put on Hegel because his monumental three–volume work "*The Science of Logic*" contains his basic philosophy. He tried to derive everything by means of his dialectic logic, which is an idealistic approach.

A quite different form of "panlogism" was proposed by the great physicist John Archibald Wheeler (Misner et al. 1973, pp. 1211–1212). The basic "building blocks" of the universe, even more fundamental than elementary particles, are *logical propositions* p, q, r, ... of the form defined in sec. 2.1 ("Logic of propositions")! This is the most radical form of panlogism, much more extreme even than Hegel's dialectic.

Later on, Wheeler modified this view by retaining only the "truth values" (1 or 0) of propositional logic (sec. 2.1). In informatics, each alternative, 1 or 0, represents a *bit* (*bi*nary uni*t*). So the world consists of "bits", or according to Wheeler (1994, p. 296): *It from Bit*. Remarkably enough, Wheeler's bits are much the same as Weizsäcker's (1985, p. 392) *Ure* (Uralternativen, basic alternatives). The same idea seems to have been found independently. (I am sure that both Weizsäcker and Wheeler are much too great personalities as to start a quarrel of priority.)

Dialectic materialism. As we have already mentioned in sec. 2.5, Marx, Engels and Lenin applied dialectics to nature itself, considered to consist of matter only. The *dialectic contradictions lie in nature* (matter) *itself*, they are in our thinking only because we think about nature. Mind is an *emergent property* of matter, it does not exist independently of matter. Engels' well–known example of a dialectic process in nature may appear somewhat farfetched: a wheat grain falls to the ground and ceases to exist as such (*negation*), giving rise to a new plant which, on dying (*negation of negation*) gives rise to the grain "at a higher level", i.e., greatly multiplied (cf. also New Testament, St. John 12:24).

Hegel's dialectic idealism appears more refined and sophisticated, but dialectic materialism (Engels' dialectic of nature) may be more concrete, "down to earth". The difference may, at least partly, be verbal rather than real: whether one calls the basic substance "spirit" or "matter" has enormous emotional significance, but is of little actual relevance if the logical structures are identical. For instance, Lenin was an enthusiastic and highly intelligent reader of Hegel. Both Hegel and the dialectic materialists claim the monist and "pantheist" thinker Spinoza as their intellectual ancestor. In fact, *monism* (*one* basic substance) seems to be what matters most.

True, Findley (1958, p. 58) speaks of "the Marxists, who try to operate Hegelian machinery with quite alien and unsuitable fuel", and the above example of the grain of wheat seems to confirm this opinion. On the other hand, Findley (1958) says also that Hegel "is more nearly a dialectic materialist than most Hegelians have realized". Again: "There is, however, as much materialism in Hegel as in Marx, since matter is for him certainly a stage in the 'Idea'. (Just as there is certainly also a strong strain of teleological idealism in the supposedly scientific materialism of Marx.)" (Findley 1958, p. 23). (The historic and economic theories of Marxism are not a subject of the present book.)

Let us look more closely to the question whether dialectics is a *property of nature* or a *property of our thinking* about nature. This question has an analogy in mathematics applied to nature. Does a "SUPERB" physical theory in the sense of Penrose (sec. 6.5), such as the general theory of relativity or quantum mechanics, reflect a property *of nature itself*, as Penrose seems to think, or is it only a law *of our thinking* about nature, according to Kant and the neopositivists? I think the answer is neither a clear "yes" nor a clear "no" (this would be another instance of Whitehead's "fallacy of misplaced concreteness") but seems to be somewhere between these extremes. My own very personal view is that a SUPERB theory of physics seems to be closer to being a property of nature, whereas a principle of dialectics seems to be more related to our thinking. At any rate, one can very well be a "dialectic dualist"! Remember also Spinoza's "*Ordo et connexio rerum idem est ac ordo et connexio idearum*" (sec. 6.5), which emphasizes the supposedly identical logical structure of our mind and of nature.

According to classical philosophy, logical structure is not all that matters. Also important is *ontological status*. Ontology (sec. 1.2) studies the kind of being or existence in philosophy. In this sense, materialism and idealism, monism and dualism are indeed different.

Pluralism. It is doubtful whether, if we decide to have more than one substance, the number 2 (dualism) is sufficient. Already Descartes and his followers had an additional fundamental substance: God.

In science, we have at least three levels: inorganic matter, life, and mind. Whether these are different "substances" or "emergent properties of matter" (as dialectic materialism would have), appears to me a matter of terminology rather than a fundamental distinction. Of essential importance is only the interaction between the various levels by upward or downward causation.

So, as a working hypothesis, I suggest that we might accept a *hypothetical pluralism*, without excluding a reduction to a "pluralistic monism" if it should become appropriate.

It is clear that all these "isms" are rather vague and by no means do justice to the many–faceted work of the great philosophers. As we have just mentioned, Hegel was in many respects a dialectic materialist rather than an idealist or "panlogist", and to dispose of Whitehead simply as a panpsychist is about as enlightening as regarding Richard Strauss as a late romantic composer.

The concept of a *set of possible solutions* considered in the context of inverse problems (see end of sec. 3.8) may also be applied to philosophy: rather than trying to find a unique philosophical system, let us look for a set of possible philosophical theories that are all compatible with the given data. In this sense we may tentatively define: *Philosophy is the set of all possible* (true, interesting) *theories* (models, perspectives) *regarding the universe.*

5.2 The three–world model of Popper and Eccles

In verity, an independent world
Created out of pure Intelligence.
William Wordsworth, The Prelude

This model was explicitly introduced by Popper and Eccles (1977) and has become rapidly popular, because it is a useful "reference model" also for those who disagree with it.

World 1 consists of all physical processes, it is the "universe of physical entities", making up our usual physical, chemical, and biological world, the "objective" world as explored and described by the natural sciences.

World 2 consists of our mental processes, sense perceptions, thoughts, ideas, emotions, pains and joys, in short, it is the world of our subjective experience.

World 3 consists of the "objective" results of man's scientific, literary, artistic, musical etc. activity. It encompasses, for instance, mathematics and logic, the theories of physics, chemistry, and biology (as found in the pertinent textbooks), Shakespeare's dramas, Goethe's poems, Beethoven's symphonies, computer programs, and, of course, all languages ever spoken by humanity. We may say, it is everything "spiritual" that transcends the mental life of a human individual. The difference between World 2 and World 3 is that between subjectivity and objectivity (or, if you prefer, intersubjectivity). It is basically identical to Plato's world of ideas, and is subject to the criticism which has been leveled against Platonism, first by Plato himself (in his dialogue "*Parmenides*") and then almost by every philosopher from Aristotle to the present time. Nevertheless, it has been indispensable at least as a level of reference. Alfred North Whitehead has characterized Western philosophy as "a set of footnotes to the work of Plato" (cf. sec. 5.4).

It is easy to criticize the terminology. World 2 and World 3 are not worlds in any geographical or astronomical sense, worlds to which you can travel on cheap tourist–class tickets.

There is, however, related terminology in general use. The term "universe of discourse" may be considered a subset of World 3, and the "realm of fancy" belongs partly to World 2 (subjective daydreaming) and partly to World 3 (fairy tales known to many people). The "World of art and letters" is also a subset of World 3.

This "world" terminology is very catchy and therefore useful, even though (or because) it is rather controversial.

Least controversial, of course, is World 1, the world of everyday life and of science. Only solipsists and some extreme idealists will deny it.

Materialists may deny the reality of World 2: mental processes, they say, are nothing but neural operations in our brain. This may be so, but nevertheless terms such as "tooth–ache", "laugh", "anger", or "noise" exist, and, in good Russellian manner, define World 2 as the set of all these terms. (Anyway, the bad headache I get when writing these lines, is for me a somewhat painful proof of the existence of World 2.)

World 3

Most controversial is the "reality" of World 3. Let us think of mathematics, considered as a prototype of this world already by Plato. Do mathematicians "invent" or "discover" their theorems? Most non–mathematicians would speak of "inventing", most mathematicians speak of discovery. Only a year ago (June 1993), there was a mathematical sensation: a proof of Fermat's Last Theorem ($x^n + y^n = z^n$ is unsolvable for integers x, y, z, n, except for $n = 2$, cf. sec. 2.3) was finally found, after Pierre de Fermat had formulated this theorem in 1637 and claimed to have found a proof but not given it. Since then, generations of mathematicians have tried in vain to find such a proof. (I just — April 1994 — learned that even the recent "proof" contains a serious flaw.) Thus, the general feeling of mathematicians is that the theorem is "already out there", just waiting for a proof (or disproof). Mathematicians are constantly busy to prove the "existence" of a solution to a complicated equation. Thus, in some way, the mathematical objects exist, just waiting for being discovered. Primarily, π is not "in the sky", but in the strange but obviously real mathematical universe (Barrow 1992).

All great mathematicians have believed in the reality of this mathematical universe, from Pythagoras to Penrose (1989). Bertrand Russell, who in his philosophy stressed empiricism, was a Platonist in logic and mathematics. Gödel emphasized the need of Platonism in logic and philosophy.

Of all contemporary sciences, mathematical physics is outstanding for the imaginative richness of its structures, combining exotic strangeness and intricate complexity with elegance and conceptual simplicity. This is almost entirely due to the use of elaborate and very abstract mathematical structures. It is difficult, if not impossible, to a "real" mathematician to deny that the strangely beautiful world of mathematics does not exist "somewhere out there", and to accept that this is merely the effect of the more or less random firings of neurons in our brains.

Obviously, prime numbers (2, 3, 5, 7, 11, ...) exist in some sense objectively, otherwise we would not have Bernhard Riemann's theorem on the distribution of prime numbers. This theorem, to me the most beautiful theorem of all mathematics, expresses the distribution of prime numbers (which are *real integers*), by means of an integral formula involving a certain analytic function of a *complex variable*, Riemann's zeta function. Of this function, we have Riemann's conjecture that all of its zeros lie on a certain vertical line in the complex plane, but this has not yet been proved. In a way, the proof of this conjecture would be much more important than the proof of Fermat's theorem mentioned above. The famous mathematician David Hilbert was asked what his first question would be if he fell asleep and woke up after 200 years. His question is reported to have been: "Has the Riemann conjecture been proved?". Where would this zeta function exist if not in World 3?

To me it is particularly impressive that Karl Popper, otherwise a sober empiricist and by no means an admirer of Plato, forcefully and (to me) convincingly argued for the independent existence of World 3: mathematics is more than the (material) set of all books on mathematics, Beethoven's 5th symphony is more than the set of its printed scores and innumerable performances and recordings up to now.

It is clear that "existence" has for World 3 a completely different meaning than for World 1 or also World 2. In a dualistic philosophy, World 1 and World 2 together constitute our "real" world in the everyday sense, whereas World 3 "exists" in a completely different way.

"Existence" is a highly ambiguous and problematic word anyway, but this word "exists" : we may look it up in any dictionary.

So even if you do not like the "existence" of World 3, consider it simply as a name for the set of all mathematical, logical, and other abstract terms, such as all terms contained in the usual dictionaries of English and of other languages.

Since *ontology* studies the various kinds of being or existence, the three worlds differ mainly in their ontological status; see also secs. 1.2 and 5.1.

At any rate, the Three–World terminology is very useful. As we have already said above, materialists may consider World 2 a subset of World 1; Spinozists may regard World 1, on the one hand, and World 2 + World 3 together, on the other hand, as the two "attributes" of the *one* divine substance; Eastern philosophers together with Schrödinger (1958) who hold that there is only *one* universal mind of which we all participate, will put World 2 ≡ World 3, Russell, in his "neutral monism", may put World 1 ≡ World 2, etc.

To the medieval philosophers, the world of ideas, our World 3, was the *mind of God*; this terminology is alluded to in the title of the book (Davies 1992). Hegel calls World 3 the "objective spirit" (not a bad terminology, after all).

Less bombastically, we may say that World 1 is governed by "fuzzy" logic and "fuzzy" mathematics (sec. 2.4), whereas World 3 is the realm of exact logic and mathematics. This point merits further elaboration, which will be done in the next subsection.

R. Bošković (1711–1787) gave an ingenious interpretation of World 3 as a mathematical space which is governed by *potentiality* in the sense of Aristotle (sec. 5.4). World 3 is the seat of the human mind, whereas matter and the human body belong to World 1 (World 2 may perhaps be considered a subset of World 3). The interaction between World 3 and World 1 explains why nature is governed by mathematical laws. Not all mathematical structures are realized in nature (not everything that is *potential*, i.e., possible, becomes *actual*, i.e., real), but theoretical physics can make free use of the inexhaustible treasure of mathematics that is contained in World 3. Furthermore, if the human mind is located in World 3, the direct access we have to mathematics according to Penrose (see below) is explained.

Another application of these ideas is the interpretation of quantum mechanics in terms of potentiality and actuality (sec. 3.5). An excellent book about Bošković is Ivan Supek: "*Rudjer Bošković*", Croatian Academy of Sciences and Fine Arts, Zagreb, 1989. More about Bošković will be found in sec. 6.6.

How is exact thinking possible?

Consider mathematical reasoning. Logical and mathematical thinking are proverbially rigorous. How can our brain perform exact thinking?

To see the problem, take any mathematical theorem about a circle, e.g., its definition: the circle is the geometrical locus of all points whose distance from a given point is constant; in other terms, the circle is a curve of constant radius.

From the times of Euclid, millions of schoolchildren have learned this and some of them even understood the definition and could work with it.

Now comes the paradox: *nobody*, not even the greatest mathematician, *has ever seen or drawn a mathematical circle*. Nobody (I really mean *nobody*) has ever seen or marked a point, and I dare say that probably nobody will ever by able to do so.

What is the reason? Logical, mathematical, and other axiomatic systems (secs. 2.1 to 2.3) are *rigorous*, that is, absolutely accurate, at least in principle. For instance, $2 + 1 = 3$ and not 2.993. Logical and mathematical objects belong to World 3. The fact that a mathematician, whose mind belongs to World 2, is able to perform a rigorous logical deduction or find a rigorous mathematical proof which is recognized as such also by his fellow mathematicians, is very remarkable indeed. Mathematicians have discovered all properties of and theorems about a circle, without ever having been able to construct one on paper.

But what about the circles constantly used in illustrations in books on geometry etc.? *They are not exact circles*, as one easily sees by looking at them with a magnifying glass or under a microscope. At best, they are "fuzzy" realizations of exact, or "real", circles!

Some mathematicians write books full of geometric theorems and proofs, which do not contain a single figure. All theorems must be derivable from the axioms by logical deduction only. It is true that most such books do contain figures, but only as an aid to better visualize the geometric situation.

Thus logicians, mathematicians etc. appear to be capable of exact thinking, of dealing with World 3 objects directly. Thus there seems to be an intimate relation between World 3 and World 2. In a way, exact circles, being objects of World 3, *can be transferred directly to World 2*.

Now comes the surprise. Circles *cannot be transferred directly to World 1*! Realizations in World 1 of abstract World 3 objects such as points, straight lines, or circles are *always* approximate only! (This holds at least for continuous objects; one might argue that integers occur in World 1 more directly: a basket containing five apples may be considered an "exact" realization of the integer "5"; cf. sec. 4.4.)

Thus we have the following scheme of objects:

in World 3:	exact,
in World 2:	exact (at least in principle),
in World 1:	fuzzy (at least in general).

This seems to be a clear indication that World 1 and World 2 are essentially different.

Analogy with computers. Now it is well known that by computers (which belong to World 1), also *exact* logical and mathematical operations can be performed. The underlying computer programs (software), however, belong to World 3. At the danger of overstretching the analogy, we may say that the program adapted so as to be able to serve as input for a particular computer, "belongs to World 2 of that computer".

Below we shall see that the computer analogy cannot be perfect. Still we may say that World 1 and World 2 are essentially different, *at least* as much as hardware and software in machine computation.

Mind vs. computer. It is claimed, however, by the most eminent mathematicians that, in a way, they have *direct access* to World 3 which goes beyond formal (algorithmic) computation or logical deduction. This is most strongly and convincingly argued in (Penrose 1989, pp. 416–423). From p. 418 we quote

> Mathematical truth is *not* something that we ascertain merely by use of an algorithm. I believe, also, that our *consciousness* is a crucial ingredient in our comprehension of mathematical truth. We must "see" the truth of a mathematical argument to be convinced of its validity. This "seeing" is the very essence of consciousness. It must be present *whenever* we directly perceive mathematical truth. When we convince ourselves of the validity of Gödel's theorem we not only "see" it, but by so doing we reveal the very non–algorithmic nature of the "seeing" process itself.

What Penrose alludes to here, is the fact that in Gödel's proof (sec. 2.3) we can "see" by informal reasoning that Gödel's basic proposition G, though algorithmically *unprovable*, is nevertheless *true*. Cf. the quotation of Findley at the beginning of sec. 2.5.

Lucas' proof of free will. Lucas (1970, §25) has used Gödel's theorem in a similar way. The essentially non–algorithmic character of human thought as exhibited by Gödel's theorem shows that our thinking cannot be the activity of a deterministic "thinking machine" because such a machine can only work algorithmically. More about this will be found in sec. 6.4.

Thinking vs. acting. For those who believe that thinking is simply an activity of the brain or the nervous system, essentially similar to its activity in directing our bodily movements, we have a surprise in stock.

In its thinking function, our brain works (or is able to work) *exactly*, as we have seen. Thus, by reasoning, we can establish rigorously valid theorems about circles.

Now let us use the brain as a "steering unit" for a bodily movement, e.g., the actual drawing (on paper or on the blackboard) of a circle. If we hope to be able to draw a circle with equal exactness, we shall be badly disappointed as we have seen above. Nobody is able to trace an exact circle, cf. Fig. 6.2 on p. 255.

The reason is, of course, the inherent "fuzziness", the inevitable random background in World 1, cf. secs. 4.4 and 6.5.

Now comes the essential question. If this random background affects our brain and nervous systems with respect to bodily movements, why does it not affect our brain equally in its thinking activity? Comparing the brain to a parallel–processing *digital* computer helps, but is only a more or less perfect analogy because of the non–algorithmic character of human thinking and because of the fact that the digital *firing* of the neurons does not yet imply that the brain is a digital *computer* (sec. 1.1).

As we have remarked at the beginning of sec. 4.4, only *integers* can be determined *exactly*, whereas *continuous* quantities can only be determined *approximately*. In computer language, this is the distinction between *digital* and *analog* computers. It is thus tempting to relate exact thinking to some digital operation of the brain; some analog (continuous) operations of the brain would then be responsible for the inexact bodily movements. There may be an element of truth in this comparison, but almost certainly matters are not that simple. At any rate, the "exactly thinking I" may be considered as a "center of command" for logical operations, which has no direct analogue in World 1 which is always "fuzzy".

To me, this is another strong argument for the independent character of World 2: this "center of command" seems to be identical to Penrose's "consciousness" (1989, pp. 409–413) and to the "self" of Popper and Eccles (1977) and to be essentially non–material in order not to be subject to random fluctuations.

5.3 Subject and object

The belief in an external world independent of
the perceiving subject is the basis of all natural science.

Albert Einstein

Philosophical problems of perception

Subject and object are basic categories of the theory of knowledge, or *epistemology*. Nothing seems to be more natural and straightforward than my (subject) seeing a tree (object) and listening to the song of a bird (object) sitting on one of the tree's branches.

Neurophysiology has shown us how complicated the problem of sense–perception is (secs. 1.3 and 1.4). Also from a philosophical point of view, however, human perception raises some fundamental problems.

(1) *Action of object on subject*. How can a tree which may be a hundred meters away, cause me, the subject, to see it? The obvious answer is that it is an empirical fact that I do see it. How this is done is explained by science: the tree (partially) reflects the light of the sun, which strikes my eye, is focussed by the eye's lense and thus produces an image on the retina. This image is then processed by the brain, as outlined in sec. 1.3. Similarly the sound waves produced by the singing bird hit my ear, are captured by it, analyzed harmonically etc.

(2) *Possibility of illusion*. Optical illusions are frequent, ordinary mirrors even more so. How do I know that the tree really exists and is at the exact spot where I see it? Or am I just dreaming? Or am I constantly dreaming, and all the surrounding world exists only in my imagination?

This is the extreme position of *solipsism*, which we have already discussed in sec. 5.1. As we have seen, it cannot be refuted. However, it cannot be proved either. From a pragmatic point of view, a consistent solipsism is not a practical way of life. If my surroundings are not real anyway, I may as well cross a busy street without paying attention to the traffic. Sooner or later I shall then dream that I am in a hospital, or my dreaming will have ceased altogether. So an honest solipsism is not an advisable strategy for survival, and this may be why any formerly existing solipsist animal or human being has been eliminated by evolution (sec. 1.4).

Descartes, starting with his "*Cogito, ergo sum*" (I think, therefore I am), from a solipsist position, argued that God, in His goodness, cannot have admitted that I am constantly deceived by illusions, so the surrounding world must be real after all. The history of philosophy has shown how many philosophers, after having argued themselves into an impossible situation, have called for God as a "*deus ex machina*" to save them, cf. also sec. 1.2. Therefore we become wary of looking at God as the Universal Metaphysical Problem Solver, and we consider Descartes' argument with some reserve.

Anyway, if we accept the existence of an external world at least as a working hypothesis (scientific realism, see sec. 5.1), we are in a position to rationally explain also dreams, optical and other illusions, abnormal perceptions, etc. It also increases our chances of survival.

(3) *Knowledge of non–knowledge*. If we do not know something, how do we know that this "something" is out there, waiting to be discovered? The answer is that our knowledge increases gradually, starting from things already known and proceeding by hopes, expectations, experiences, predictions, working hypotheses, etc. When, driving a car, we see a ball rolling across the street, we watch out for the child following it and step on the brake. As Whitehead has said, we should not even see an elephant unless we expect some animal of this kind. A person may sit in a room, complaining about the terrible noise to which he is exposed, and may not notice that it is Beethoven's Seventh Symphony. The bird lover walking in a forest recognizes dozens of birds whereas the accompanying person does not notice anything.

(4) A related problem is *knowledge a priori and knowledge a posteriori*. According to Kant, logic (analytic) and mathematics (synthetic a priori) are tools by which we explore our surroundings. Well–established theories are also knowledge a priori. Kant regarded Euclidean geometry, too, as absolutely true a priori knowledge. General relativity, however, has taught us that this is not warranted in an absolute sense, though to an extremely good approximation. So our present theories of physics are considered excellent a priori knowledge, but always subject to replacement by a better theory or by a more appropriate paradigm (sec. 3.10); see also Quine's model at the end of sec. 4.5.

(5) *The problem of truth*. How do we know that our knowledge corresponds to reality? One answer is: by testing it through observations and experiments (verification and falsification). Such an experimental verification can never be absolute, not only because of experimental errors (sec. 4.4). For instance, take as law our standard example: every day, the Sun rises at a precisely predictable time. It may fail to do so tomorrow because, during the preceding night, the

Sun may have exploded or the Earth destroyed by a giant comet. Not even a falsification necessarily proves a law wrong: the experimental setup may have been inappropriate or measuring errors may have provided the false impression that the theory was incorrect. This is the problem of *induction* already discussed at some length in sec. 3.9.

All this is relatively trivial. A deeper question is whether reality really "is" as it looks. For instance, is space "really" three–dimensional? It may be five–dimensional or ten–dimensional, only we may not be capable of perceiving more than three dimensions. Are there aspects of reality which are inaccessible to science? Are there Kantian "things–in–themselves" which are unknown as a matter of principle (p. 230)?

As a partial answer, let us consider the spectrum of electromagnetic waves, with wave lengths from several kilometers down to 10^{-13} m. The visible spectrum, from red down to violet, is on the order of 5×10^{-7} m, above there are infrared and longer electromagnetic waves; below there are ultraviolet waves, X–rays and gamma–rays. We know very well and use also electromagnetic waves above and below the visible part of the spectrum: they are accessible to indirect physical observation.

So it seems that we know practically everything which in principle can be observed by present–day physical, chemical, etc. experiments. Other things, e.g., some elementary particles in physics, have been predicted by theory but not yet been discovered (this is a nice illustration of "knowledge of non–knowledge" mentioned above). Also higher dimensions, which are inaccessible to direct observation, are mathematically fully understood (sec. 2.6) and are under consideration as candidates for certain "unified theories". Infinite–dimensional spaces are, of course, standard tools in quantum mechanics (without implying that "ordinary" space is more than three–dimensional). (See, however, also Bohm's (1980) "enfolded reality" and "implicate order", cf. sec. 3.5.)

It would be very unwise indeed to exclude that there are physical or other phenomena which are completely different from all that we know today and which cannot be observed by contemporary science, and *for which we do not even know where to look*. As an example, think of quantum phenomena before 1900.

What we can say is that our knowledge of the physical world is essentially *correct* in corresponding to some external reality; this does not exclude the possibility or even probability that it is *incomplete*.

Think of trying to catch a fly. Why is it so difficult? The fly's knowledge of the external world is certainly extremely rudimentary: it probably comprises mainly the (instinctive) knowledge of how to get food, how to escape enemies, etc. This limited knowledge, however, is nearly perfect. Perceiving a hand (not necessarily as a hand, but as a dangerous object) approaching rapidly, the fly escapes even much more rapidly in *exactly the right direction*! This ability was certainly acquired through evolution: only the "most intelligent" flies survived. Within its very narrow limits, the fly's knowledge of the external world is perfectly "true": it is sufficient for performing exactly the right reactions.

So one answer why, in some way, man's knowledge is "true", has been given: only those human beings who have a reasonably correct relation to their environment, have been able to survive evolution. This biological consideration cannot be expected to provide more than a partial answer to the complex question of "truth", but it seems to play a certain basic role.

One of the best philosophical treatises on the theory of knowledge is (Hartmann 1965), see in particular Chapter 6.

Monadic structures

Fig. 5.3 schematically illustrates a "monadic structure", consisting of a subject A and an object B. It is probably the simplest and most fundamental structure in epistemology, relating an object, considered to be at a location B, to a subject A (represented by the eye of the observer). In order to avoid Whitehead's "fallacy of simple location" (sec. 5.1), we consider the object with respect to a certain location B without excluding the possibility that the object is implicitly present also at other locations, cf. Fig. 5.1 on p. 200.

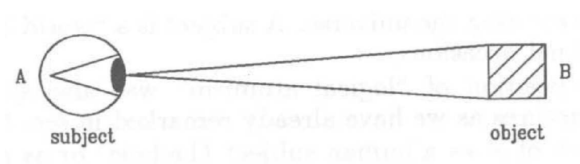


Figure 5.3: A monadic structure

The concept of *monad* was introduced by the great mathematician and philosopher Gottfried Wilhelm von Leibniz (1646–1716). A monad is essentially an unextended "pointlike" subject A mirroring the whole universe, that is, A as the center of perspective for all objects B. Thus the monad is considered a pointlike "subject" together with all objects B as seen from A. Symbolically we may represent a monad perhaps somewhat as in Fig. 5.4.

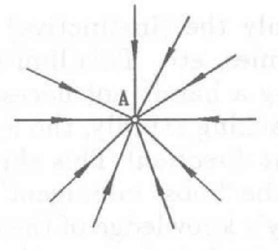


Figure 5.4: A monad A

According to Leibniz but in modern terminology, there is a monad situated at every point of space. Human persons (souls) are particularly well–developed monads, and God is the greatest monad of all. All monads are "windowless", that is, closed to each other and without interaction with each other except with God. By some kind of "pre–established harmony", all monads are kept in tune with each other by God, like perfectly synchronized clocks.

Thus, according to Leibniz, monads are non–interacting "substances", reminding of the two non–interacting substances of Descartes: matter and mind, which may also be kept synchronized by some pre–established harmony (sec. 5.1).

Just as Popper and Eccles have allowed these substances to interact, Alfred North Whitehead has worked out a theory of monads (he calls them "actual occasions") which, in fact, do interact.

Whereas Leibniz' souls are more or less permanent, Whitehead's "actual occasions" last only for a moment, they are, so to speak, space–time points mirroring the universe. A subject is a "world line" (sec. 3.7) of related actual occasions.

Russell's position of "logical atomism" was also sympathetic to monadic structures as we have already remarked in sec. 5.1.

If we think of A as a human subject (Leibniz) or as a point of the world–line of a human subject (Russell, Whitehead), this looks reasonable enough. If, however, we regard, with Leibniz, all points in space (or all material atoms) or, with Whitehead, all points in space–time (all actual entities) as monadic subjects with some kind of primitive "soul" (Whitehead speaks of "mental pole"), we get a doctrine already known to us as *panpsychism*: everything (Greek: *pan*) has a soul (Greek *psyche*), which seems hard to accept for many people.

Let us not forget, however, that Leibniz and Whitehead belong to the greatest mathematicians, logicians and philosophers of their time.

If they held a doctrine which apparently contradicts common sense, they had good reasons as everyone can see on reading Whitehead's books.

It is true that many philosophers use to criticize and even ridicule their colleagues much more than we are used from other disciplines, but we outsiders should exercise great restraint in order to avoid too cheap criticisms. In fact, modern physics tells us one lesson: many theories may be very far from "common sense" and still true, even unavoidable: think of special and general relativity, quantum theory, and modern elementary particle physics, which are all extremely abstract and "counterintuitive". In comparison with these theories, Whitehead's "actual entities" are models of simplicity and common sense.

Again, Niels Bohr's dictum comes to our minds: "Your idea is, of course, crazy. The problem is only whether it is sufficiently crazy to be true."

According to panpsychism, "mind" is simply "matter" *seen from within*, cf. (Teilhard de Chardin 1955) and the article by Globus in (Globus et al. 1976).

Even quite apart from the question of panpsychism, monadic structures are important in relativity where A corresponds to a particular reference system, and in quantum theory where A is the observer in the standard Copenhagen interpretation; cf. secs. 3.4 and 3.5.

For the mathematical reader, this is very clearly expressed by the simple matrix formula, well known from quantum theory,

$$l = \mathbf{e}^T \mathbf{L}\,\mathbf{e} \tag{5.2}$$

where, in the observed average value l, the effect of the object (physical quantity) is the matrix $\mathbf{L}$, and the effect of the observer is the unit vector $\mathbf{e}$ ("state vector"); $\mathbf{e}^T$ denotes the transpose of $\mathbf{e}$. (For an even more mathematical reader we shall not conceal the fact that the "observable" $\mathbf{L}$ represents an *infinite* matrix in the Heisenberg sense or, equivalently, a linear operator in the Schrödinger sense; the vector $\mathbf{e}$ is an infinite "state vector" or a "state function".) The measured value l thus is the projection of the "physical quantity" $\mathbf{L}$ onto the reference system $\mathbf{e}$ of the observer. For more details see sec. 3.5.

Priority of subject or object?

This question is somewhat like the problem: which comes first, hen or egg?

Philosophy quite naturally starts with the subject, that is, with *idealism*. Equally naturally, science starts with the object, that is with *realism* or *materialism*.

As we have already remarked in sec. 5.1, realism is not absolutely identical with materialism. Realism only emphasizes the priority and existence of the external world, but does not deny that other persons may have minds, as materialism does (as a matter of fact, materialism also denies the reality of mind in the subject).

In sec. 5.1 we have seen that scientific realism, in the form of the mathematical laws of physics, contains "non–material" objects of World 3. This synthesis of "realism" and "idealism" was seen to make "matter less material and mind less mental" (Russell).

In fact, like in the hen–egg problem, subject and object are inseparably connected *in human knowledge*, as already Plato, especially in his dialogue Parmenides, has emphasized (e.g. Speiser 1952, p. 13): no subject without object, no object without subject.

In science, this connection is particularly strong in quantum theory, and very weak in sciences like paleontology. There the scientific subject may come millions of years later: nobody will question that dinosaurs have existed even without human observers watching them with awe and fear, but nobody will question either that human observers are necessary for the *science* of paleontology. Nature may very well exist without human observers, but philosophy certainly not, and not even science.

Fichte's iteration

Matters become difficult in introspection: the mind is subject and object at the same time: *the thinking thinks the thinking*. This has already been mentioned in sec. 2.5; see the quotation of P.M. Møller which was so dear to Niels Bohr, and the subsequent closely similar quotation from Fichte. In fact, this does lead to an infinite regress of "I's" but that may not really matter: there are also many kinds of mathematical infinities, so that World 3 (sec. 5.2) is certainly not subject to overpopulation as our small Earth is.

Still, there are the well–known paradoxes of self–reference which we have already discussed in secs. 2.3 and 2.5. As Sir John Kendrew put it (Duncan and Weston–Smith 1977, p. 207):

> Perhaps the most fundamental of all the difficulties encountered in biological research is that the investigator cannot detach himself from the system under study because he himself forms part of that system ... in psychology the investigator is himself one of the subjects under study.

I believe that an iterative process given by Fichte in his Wissenschaftslehre of 1804 provides, if not a solution, at least an orderly approach; cf. also sec. 5.1. Denoting realism by R and idealism by I, we have an iterative approach which may be represented as a sequence of syntheses in Fig. 5.5 (similar to Fig. 2.9 on p. 48). The first iteration has already been discussed in sec. 5.1: R_0 is naive realism, I_0 is scientific idealism, R_1 is critical or scientific realism.

In his Wissenschaftslehre of 1804, Fichte has given 10 of such syntheses, according to M. Guéroult ("*L'évolution et la structure de la doctrine de la Science chez Fichte*", vol. 2, pp. 136–137, Strasbourg 1930). We may even continue this process *ad infinitum*, in order to arrive at what philosophers call the *Absolute*. Mysticists claim to have direct access to the Absolute by meditation, rationalists like Fichte and Hegel ("Wissenschaft der Logik", 1812) prefer an iterative approach. Will the iteration of Fig. 5.5 converge? No, in the sense that it does not converge to any ultimate "ism", in much the same way as the sequence (2.18) on p. 55 does not converge to a *rational* number. Just as Cantor *defined* the *irrational* number $\sqrt{2}$ as the sequence (2.18), we *define* with Fichte and Hegel the Absolute by the complete infinite sequence of syntheses of Fig. 5.5.

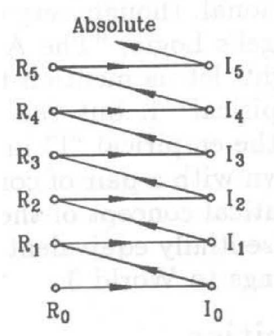


Figure 5.5: Fichte's iteration

Fichte's and Hegel's works are of "Rabbi type 3" (see Preface), which means dark, difficult and fascinating. For a clear description of this process (Rabbi Type 2) cf. (Speiser 1959, Chapter II: Fichte's Wissenschaftslehre von 1804). At this point I should like to express my appreciation for the books of Andreas Speiser who treated these difficult problems as a lucid mathematician, and my thanks to the well–known physicist Hans–Jürgen Treder (Potsdam–Babelsberg) with whom I had brief but wonderfully enlightening discussions on Parmenides, Plato, Fichte, Speiser, and unified theories of physics.

Plato, in his dialogue "*Parmenides*" *starts* from the Absolute, considered as the One, and tries to get *down* in the opposite direction of Fig. 5.5, which is much more difficult to understand in spite of Speiser's (1959) brilliant commentary.

At this point I should like to mention Weizsäcker (1971, Chapter IV, secs. 5 and 6) who relates Plato's "*Parmenides*" to biology and quantum theory in a most instructive way.

Starting with (Schrödinger 1958), relations between quantum physics and Eastern mystical philosophy have become quite numerous (Bohm 1980; Capra 1976; Moser 1989). This may be understood also as a protest against the dominating influence of logical positivism (sec. 5.4), to restore the balance between "atomistic" and "holistic" thinking.

Fichte and Hegel provide a rational approach which is accessible also to those who (like the present author) have not found a direct access to mystical meditation.

As an ultimate synthesis, the differences between realism and idealism vanish in the Absolute. Subject and object become One, which Buddhists and other mystics try to achieve by meditation. An incomparable description in rational, though very difficult, language is found in the final section of Hegel's Logic, "The Absolute Idea".

To put matters straight, let us mention that Fichte's "I" is not his personal individual "empirical" I, but the counterpart of the "I" in World 3. It is related to the empirical "I" in precisely the same way as an empirical "circle" drawn with a pair of compasses on the blackboard is related to the mathematical concept of the circle (a World 3 object).

Fichte's ideal "I" is essentially equivalent to Kant's "transcendental subject", which also belongs to World 3.

Logical singularities

Fichte's "Absolute" is clearly a logical singularity, as Fig. 5.5 shows. This is a general situation, which is very well described by the quotation from John Kendrew given above.

Let us recall some examples of such a subject–object interference:

(1) *Mind*: Thinking about thinking (Plotinus, Fichte, Hegel; just discussed).

(2) *Mind*: Paradox of the liar, Gödel's theorem (sec. 2.3).

(3) *Mind*: Paradox of introspection (introspection may disturb the mental phenomenon under investigation: "blue elephant", sec. 2.5).

(4) *Observer* ↔ *nature*: observation may disturb the quantum phenomenon under investigation (Heisenberg's uncertainty relation, sec. 3.5).

(5) *Observer* ↔ *life* (Bohr's paradoxon): a detailed physical investigation of a living organism may destroy it (sec. 4.5).

(6) *Observation of human test persons*: the behavior of a test person may be changed by the fact that the person is aware of being tested. This is a psychological or medical (placebo effect!) phenomenon not unlike Heisenberg's uncertainty relation (see Example 4): observation disturbs the event under investigation. Introspection (Example 3) is a special case.

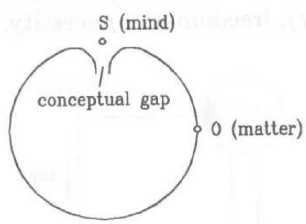


Figure 5.6: Singularity $O \rightarrow S$

A rather general principle is represented in Fig. 5.6. If the object O approaches the subject S, so as to finally coincide with it, there is a singularity. A conceptual gap has to be bridged, which is almost impossible. Here it is appropriate to continue the quotation from John Kendrew (Duncan and Weston–Smith 1977, p. 207):

At a more fundamental level the problem of the relationship between mind and matter and of the nature of consciousness seems impossible of solution because the investigator himself is conscious mind and there is a complete conceptual gap between that mind and the physical objects in which it resides.

Already Kant recognized that it is *impossible to regard the universe as an object* to be studied, because the observing subject is necessarily part of the universe. The way in which Fichte and Hegel approach the problem may, as a variation of Fig. 5.5, also be described as a spiral (Fig. 5.7), corresponding to looking at the singularity S in Fig. 5.6 *from above*. Fig. 5.7 clearly corresponds to the dialectic spiral of Figures 2.10 or better 2.12 on p. 53, and hence to Weizsäcker's "Kreisgang".

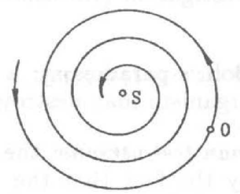


Figure 5.7: The singularity is approached by a spiral

According to Kant, considering the infinite universe as an object leads to *antinomies* (or paradoxes) regarding the beginning of the universe ($t = 0$ or $\tau \to -\infty$, cf. sec. 3.7), the divisibility of matter (discreteness vs. continuity), freedom and necessity, and the creation of the world.

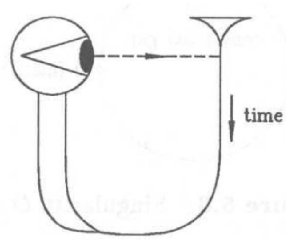
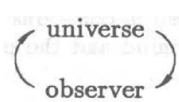


Figure 5.8: Wheeler's self–reference universe

According to Wheeler (in Duncan and Weston–Smith 1977, pp. 31–32), we have a "self–reference universe" (Fig. 5.8). The evolution of the universe has finally given rise to man, whose observation and activity give meaning to the universe. We thus have a "self–exciting circuit" (the name comes from electromagnetic dynamo theory!):



So to speak, the observer and the universe "bootstrap" each other into existence. (In this light, Wheeler's admiration (p. 180) for Eigen's "life–machine" (Fig. 4.10 on p. 179) is not surprising since Eigen's "hypercycle" is also a catalytically "self–exciting circuit".)

According to Bohr (1934, last sentence): (Let us not forget) "the old truth that we are both onlookers and actors in the great drama of existence". A similar idea has been expressed by Maturana and Varela (1987).

5.4 Historical landmarks

The safest characterization of the European philosophical tradition is that it consists of a series of footnotes to Plato.

Alfred North Whitehead

Philosophy cannot be understood without some knowledge of its historical development. A comprehensive history of philosophy is the 9–volume work of Copleston (1946). Extremely profound, sympathetic, and readable is (Jaspers 1962–1993). At this point we also mention the wonderful booklet "*Introduction to Philosophy*" by the same author (Jaspers 1953) which, together with (Bocheński 1959) and (Russell 1912) forms a triad that furnishes an ideal, easy and yet comprehensive, introduction to general philosophical thinking in all directions.

A relatively short and eminently readable (Rabbi Type 1) history is (Russell 1945). This brilliant work is, however, somewhat biased by its author's strongly expressed viewpoint of logical positivism. Especially German idealism (Fichte, Schelling, Hegel) is grossly misrepresented. On the other hand, we have excellent monographs on German idealism: (Gulyga 1990) and especially (Hartmann 1960). A general concise history of philosophy, particularly for mathematicians and physicists, is (Kuznecov 1981). An excellent anthology of texts of various philosophers is (Perry and Bratman 1993).

With (Russell 1945) we restrict ourselves to *Western* philosophy. It would be tempting to include Eastern (Indian, Chinese) philosophers, especially with a view to dialectics, but my ignorance does not permit it; see, however, (Jaspers 1962–1993).

Greek philosophers of nature

We know names such as *Thales*, *Anaximander*, and *Anaximenes* (around 560 B.C.), belonging to the Milesian school (city of Miletus in Asia Minor). Their activity would today be considered some rudimentary form of physics. They were concerned about elementary constituents of matter: water, air, earth, fire.

More important for the present view of science is *Pythagoras* (around 530 B.C.), the founder of mystical mathematics, whose ideas on the harmony of the world influenced Plato as well as Johannes Kepler. They even influenced the modern theory of oscillations including light and matter waves, fundamental in quantum theory.

Parmenides (around 480 B.C.) was the first monist philosopher: all is One, change is an illusion. He thus was the first representative of a *block universe* (secs. 3.7 and 6.3). He exerted a strong influence on Plato and on modern monists such as Spinoza and Hegel.

Heraclitus (around 500 B.C.) represented an opposite point of view: *panta rhei*, all is flowing. It is impossible to step into the same river twice. His universe is dynamic, and he may be considered the first dialectic philosopher: he discovered the dialectics of identity and difference (sec. 2.5).

Democritus (around 420 B.C.) was the first atomist. All matter consists of atoms. This hypothesis had to wait for its experimental confirmation until Rutherford around 1900 (so Democritus would never have got the Nobel Prize).

Many concepts of modern physics and philosophy were foreshadowed by these philosophers between 600 and 400 B.C. No wonder that they have exerted a considerable attraction to modern atomic physicists such as Heisenberg and Schrödinger (1954) (who did get the Nobel Prize).

Plato

Plato (428–347 B.C.) was the founder of modern philosophy and is regarded by many as the greatest philosopher of all times. According to Whitehead (1929), the European philosophical tradition "consists of a series of footnotes to Plato" (see the motto of the present section). He is not a systematic philosopher and provides questions rather than answers. He gave the theory of ideas (Popper's World 3, cf. sec. 5.2) and emphasized the importance of mathematics. He raised almost all contemporary problems of philosophy (mind and matter, God, dialectics, etc.). Rather than constructing one philosophical "system", Plato wrote dialogues of high literary rank, which permitted him to regard a problem from various angles. Thus he is the first representative of philosophical *pluralism*, which provided an inspiration to many different philosophical directions.

Aristotle

Aristotle (384–322 B.C.) was a student of Plato. He is the creator of logic and systematic metaphysics. Substance consists of *matter* and *form*. There are 4 causes:

- *causa efficiens*: causality in the modern sense (determinism), physics;
- *causa finalis*: final cause, "downward causation", biology (cf. sec. 4.1);
- *causa formalisis* (the plan of a house); and
- *causa materialis* (matter, e.g. bricks of which the house is built).

The reader may wonder about these Latin names: they come from medieval philosophy, especially St. Thomas Aquinas (see below). Of course, one may also speak of "efficient cause", "final cause", "formal cause", and "material cause".

The last two "causes" correspond to matter and form mentioned above. In contrast to Plato, his emphasis is on classification and description rather than on mathematics. He was much more interested in concrete science (physics, biology) and much more systematic than Plato who was primarily a mathematical thinker. God, to him, is the "prime mover", the first cause. Aristotle was very influential in the Middle Ages, where he was regarded as an authority in science as well as in philosophy. The very apparent completeness of his scientific work made it obsolete as soon as new discoveries were made which did not fit into his system. His main importance in science today is his creation of systematic logic, which was transcended only by symbolic logic (Peano, Frege, Russell, Gödel etc.). Through the mathematization of modern science, Plato won over Aristotle. With some pointed exaggeration we may say that Plato is the philosopher of mathematical physics, whereas Aristotle is the philosopher of biology.

Nevertheless, his distinction between *potentiality* and *actuality* is important for the difficult problem of the relation between mathematics and the physical world (sec. 5.2, p. 210), for the interpretation of quantum mechanics (sec. 3.5, p. 108), etc.

Neoplatonists

Plotinus (204–270 A.D.) took up and developed further the mystical and mathematical aspects of Plato's thoughts. His works are literary masterpieces. Goethe studied Greek in order to read Plotinus. He was not a Christian but exerted great influence on St. Augustine. His statement "the thinking thinks the thinking" may have served as a model for Augustine's theory of the Holy Trinity.

St. Augustine (354–430) is most important for our purpose because of his *theory of time* ("What is time? If nobody asks me, I know, but if I want to explain it to someone, then I do not know.") What did God do in the time before the creation of the universe? A humorous answer is: He prepared hell for people asking such foolish questions. Augustine's correct answer was: the question is meaningless because before creation (the "big bang") *there was no time*.

Boethius (480–524) was a high government official. Sentenced to death for alleged treason, he wrote in prison his famous work "*Consolationes philosophiae*" (Consolations of Philosophy). Here he makes the first allusion to the problem of a *block universe* (secs. 3.7 and 6.3). If God is outside of time, then He has the whole universe (today we would say: the entire space–time continuum) before Himself. Thus He also knows the future. If the future is thus determined, where is the place for man's free will? Boethius answers that God's *foreknowledge* (in His ever–present eternity) does not mean *predestination*, so that man remains free and responsible for his actions. For St. Augustine's and Boethius' views of time cf. also (Lucas 1973).

Medieval philosophy

The main miracle of Western philosophy in the dark time between 500 and 800 is its sheer survival: all the principal works of classical antiquity have been preserved. With *Johannes Scotus Erigena* (around 810–877) original work in the Platonist tradition started again, and with *Anselm of Canterbury* (1033–1109) it reached a first culmination.

With Anselm, the school of medieval scholastic started. It flourished in Paris and at other universities, and it is a school in which theological and philosophical questions were discussed in a professional way, based on more or less generally accepted premises; sometimes also these premises were subjected to discussion. This kind of school philosophy has contemporary analogues in the schools of Marxist philosophy in the communist countries, and of logical positivism in the West.

The most eminent scholastic philosopher was *St. Thomas Aquinas* (1225–1274). Whereas his predecessors were Platonists, Thomas used the work of Aristotle as the basis of his philosophy. This was his strength and his weakness. He gained a rigorous and consistent system, but he partly lost Plato's inspiration. Thomas' merit was also a strict separation between philosophy and theology. Until recently, he was the official philosopher of the Roman Catholic Church.

Scholastic philosophy continued its systematic and detailed work. Well–known is *William of Occam* (around 1280–1350) because of the principle of *Occams razor*: "*Entia non sunt multiplicanda praeter necessitatem*": no unnecessary concepts should be introduced.

The last medieval philosopher, *Nicolaus Cusanus* (1401–1464), however, was an individualist. He was a typical dialectic philosopher. With his principle of "*coincidentia oppositorum*", the coincidence of logical opposites, he was a forerunner of Hegel and Bohr (cf. sec. 2.5).

The non–Christian (Arabic, Jewish) medieval philosophers Averroes, Avicenna, Maimonides etc. should at least be mentioned.

Descartes

René Descartes (Latin: Cartesius, 1596–1650) was the first modern philosopher and mathematician. He invented analytical geometry (Cartesian coordinates!). He started with systematic doubt about everything. His principle "*Cogito, ergo sum*" (I think, therefore I am) allowed him first to prove his own existence. (This is a typical argument of *dialectic reversal*: original doubt gives rise to higher–level certainty! Cf. Gödel's argument in sec. 2.5.) Descartes might have ended as a solipsist, were it not by an appeal to God. In fact, theoretically, what I experience and what looks real to me, might well be only a dream. However, God, being the absolute Truth, will not permit that I am deceived all the time in this way. So the external world must be real.

Descartes maintained a strict difference between matter and mind, considering them as two completely separate "substances", whose interaction is difficult to explain without recourse to God as the "Universal Metaphysical Problem Solver". His *dualism* is thus subject to Whitehead's "fallacy of misplaced concreteness"; nevertheless it has been very influential in philosophical thinking until today.

Spinoza

Baruch de Spinoza (1632–1677) is a *monist*, recognizing only *one* substance, called *nature* or *God* ("*deus sive natura*"). Matter and mind are "attributes" or "modi" of this divine substance. This doctrine is called *pantheism* (Greek: everything is God).

Spinoza's *one substance* can easily be identified either with *mind* or with *matter*, so he is claimed by idealists like Hegel, as well as by dialectical materialists.

The pantheistic, religious, side of Spinoza has attracted Goethe as well as Einstein, both of whom liked to see God in nature.

We may say that Spinoza is halfway between Parmenides and Hegel.

Leibniz

Gottfried Wilhelm von Leibniz (1646–1716) was also a mathematician and logician, besides being a great philosopher. Together with Isaak Newton, he founded the differential and integral calculus. As a logician, he wanted to invent symbolic logic, but fell short of his aim. He also constructed a computing machine.

Leibniz' main philosophical merit is the theory of *monads*, already discussed in sec. 5.3, cf. Figures 5.3 and 5.4. All monads are "windowless" and interact only with the highest monad, God.

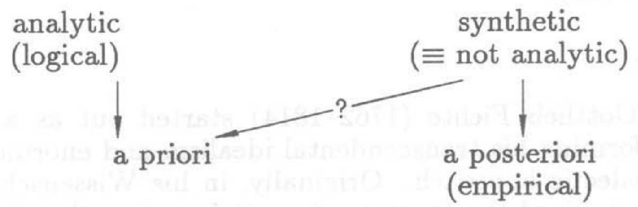
The monadic theory, artificial as it looks, has nevertheless influenced such diverse philosophers as Bertrand Russell and Alfred North Whitehead. Whitehead's "actual occasions" are interacting monads in space–time; his theory has proved surprisingly fruitful in quantum theory (Bohm 1980; Stapp 1993).

Kant

Immanuel Kant (1724–1804) is considered the greatest philosopher after Plato (with the possible exception of Aristotle or Hegel). Important and influential has been his criticism of classical metaphysics, which studies the existence of God, the immortality of human soul, the freedom of the will, and the existence of the external world in the tradition from Descartes to Leibniz. Kant admits only empirical knowledge. (According to this criterion, also Kant's philosophy, being nonempirical, is metaphysical and, strictly speaking, should also be rejected! This remark, however, should not be taken too seriously: this would be a rather gross oversimplification.)

The "*things–in–themselves*" (Dinge an sich) of the external world are unknown as a matter of principle. What exists is only a generalized subject, the "transcendental I". Kant is therefore a "transcendental idealist". Human knowledge can be *a priori* and *a posteriori*. *A priori* are *logical* identities (e.g. a man is a human person); they are also called *analytic*. *A posteriori* knowledge is *empirical*, obtained through the senses; it is also *synthetic*, which means something beyond analytical tautologies (logical identities).

Kant now asks: Is also *synthetic a priori* knowledge possible, that is, is there non–empirical and logically non–trivial knowledge? We thus have the following scheme:



Kant considers *synthetic a priori* our knowledge of mathematics, of three–dimensional Euclidean space, and of Newtonian mechanics (causality). By means of a very original argument he says that this synthetic a priori is not a property of nature, but reflects the *structure of our mind*. This "*Copernican revolution of philosophy*" (turning from object to subject) makes our mind responsible for our basic knowledge of mathematics, Euclidean geometry, and causality, which, being imposed by ourselves, is by definition *absolutely exact*. This is best understood by Eddington's example given in sec. 1.4.

The presently prevailing view is that mathematics is analytic (reducible to logic), and that the present fundamental laws of nature indeed reflect the properties of our mind, *together with* real properties of nature. Our theories are *perspectives* from which we look at nature (*theoria* = insight, a way of rega rding). A theory thus may be compared to a telescope for looking at stars, or to a microscope by which we study organic tissue. Just as a telescope or microscope can be more or less precise, even out of focus, a theory can be more or less exact (sec. 6.5).

Thus, with Eddington (1939) and Bohm (1980) we may accept *synthetic a priori* knowledge, but in contrast to Kant, this knowledge is not absolutely accurate. In fact, general relativity implies curvature of space–time through gravitation, hence deviations from Euclidean geometry; and quantum theory predicts deviations from causality. *Evolutionary epistemology* provides a contemporary biological interpretation of Kant's synthetic a priori; cf. sec. 1.4.

Modern science (relativity and quantum theory) has frequently, though critically, evoked Kant. Logical positivism has been particularly sympathetic to Kant's rejection of metaphysics.

Kant's dialectics, concerning the antinomies connected with classical metaphysics, on the other hand, has strongly influenced Fichte, Schelling, and Hegel, who, in this way, developed a metaphysics of their own. In fact, metaphysics has steadfastly refused to die, as philosophers such as Whitehead (1929) and Hartmann (1965) show; cf. also

(Stegmüller 1969).

Fichte

Johann Gottlieb Fichte (1762–1814) started out as a student of Kant, transforming his transcendental idealism and enormously developing his dialectic approach. Originally, in his Wissenschaftslehre of 1794, he emphasized the (transcendental) I, posing the dialectic triad I (subject, thesis), non–I (object, antithesis) together with their synthesis. This obviously led Russell (1945, p. 718) to write that Fichte "carried subjectivism to a point which seems almost to involve a kind of insanity". This is a classical example of a gross misjudgment of a philosopher by a colleague; such misjudgments unfortunately occur rather frequently in the history of philosophy (and it would be presumptuous to claim that the present book is free of them).

Later, in the Wissenschaftslehre of 1804, Fichte started with the object (Fichte's iteration, sec. 5.3). So, if Fichte is called an idealist, by the same token he could be called a realist. This is another example of a profound truth in the sense of Niels Bohr (sec. 2.5). The mathematician Speiser (1959) relates Parmenides, Plato, and Fichte in a remarkable "synthesis".

Hegel

Georg Wilhelm Friedrich Hegel (1770–1831) is considered the greatest dialectic philosopher. We have already discussed his dialectics to considerable length in sec. 2.5. Although he is usually regarded as an idealist philosopher, his thinking has become basic also for dialectic materialism: Karl Marx (1818– 1883) and Friedrich Engels (1820–1895). Vladimir Ilich Lenin (1870–1924), the Russian revolutionary, studied Hegel in great detail.

Needless to say, philosophers like Kant and especially Fichte and Hegel are very difficult to read. Nevertheless, Fichte's Wissenschaftslehre of 1804 and Hegel's logic, which provide two alternative ways of "ascent to the Absolute", are extremely rewarding and well worth the reading effort, just as the Rabbi's third speech mentioned in the Preface. (It is preferable, however, to start with the secondary literature, e.g. (Findley 1958).)

Logical positivism and analytical philosophy

Logical positivism, also called *neopositivism*, started at Vienna, inspired by the physicist Ernst Mach (1836–1916), with Ludwig Wittgenstein (1889–1951), Rudolf Carnap (1891–1970), and others, and by

Bertrand Russell (1872–1970) in England, where it soon became the dominant philosophy; also Wittgenstein emigrated to Cambridge in 1929.

It is based on mathematical logic (secs. 2.1 and 2.2) and on the empirism of the British philosophers John Locke (1632–1704) and David Hume (1711–1776), who had also influenced Kant, as well as on the developments of modern science, cf. (Russell 1945). It basically says that all knowledge seriously so–called consists of logic, mathematics, and empirical science.

It has become the leading philosophy of science, analyzing basic concepts such as axiomatics, deduction and induction, verification and falsification, etc. In this respect, its work is recognized by probably all serious philosophers. What characterizes logical positivism, in particular, is the claim that philosophy *is nothing else than* logical analysis of science and of language, that metaphysics is meaningless etc. This claim, of course, has been contested, cf. (Stegmüller 1969) and any book on contemporary philosophy.

Its underlying "ontology" is *logical atomism* (sec. 2.1). Besides metaphysics, it also rejects dialectics etc.

After 1945, analytically–minded philosophers in Great Britain and elsewhere recognized that they didn't have to be logical positivists. Thus they are called *analytical philosophers*; cf. (Cohen 1986).

Important philosophers who initially adhered to neopositivism later went their own ways:

Kurt Gödel (1906–1978), perhaps the greatest logician after Aristotle, emigrated to Princeton (U.S.A.) and worked exclusively on foundations of logic (apart from a pioneering paper in general relativity, cf. (Hawking and Ellis 1973)). In his later years, he was an idealist of Parmenidean type and a defender of a static "block universe".

Alfred North Whitehead (1861–1947) was a British mathematician who collaborated with Bertrand Russell in writing the monumental logical work "*Principia Mathematica*" (1910–1913). In 1924 he went to Harvard University (U.S.A.) and became one of the greatest philosophers and metaphysicians of our century. In the present book, we have encountered him on many occasions.

Karl Raimund Popper (1902–1994) originally was close to the Viennese school of logical positivism, together with Carnap, Gödel, Reichenbach, and Wittgenstein. He has written basic books on many questions of philosophy of science; cf. (Popper 1977, 1979). In his later years he critically approached metaphysics (Popper and Eccles 1977); cf. also sec. 5.2.

Ludwig Wittgenstein (see above) later on became a linguistic philosopher, concerned with language and its limitations. Like Gödel, he was a strange but enormously respected and influential philosopher, although his later philosophy was very remote from science. Wittgenstein's later thinking is expressed in short aphoristic, apparently simple but very profound statements which are often paradoxical and subtly dialectical. The great physicist Hawking (1988, pp. 174–175) writes:

> Philosophers reduced the scope of their inquiries so much that Wittgenstein, the most famous philosopher of this century, said: "The sole remaining task for philosophy is the analysis of language." What a comedown from the great tradition of philosophy from Aristotle to Kant!

I have quoted this passage not because I agree with it, but because it shows how separate great science has become from great philosophy in our times.

In fact, none of the great physicists of our century was a logical positivist, rejecting metaphysics. Einstein was a follower of Spinoza (pantheism), Wolfgang Pauli (the famous "Pauli exclusion principle" in quantum mechanics) was a mystical thinker, Schrödinger was an idealistic monist of Eastern type (all consciousness is one), and Heisenberg and Weizsäcker continue the classical tradition of Western philosophy. Finally, Niels Bohr was a typical dialectic thinker: his principle of complementarity is for a physicist what dialectics is to a philosopher. (It is very interesting in this respect to compare the physicist O. Hittmayr and the philosopher E. Heintel; cf. (Heintel 1990, p. 232–234)).

Why do scientists need classical philosophy?

In mathematics and science, the names of great discoverers are mentioned but it is very rare that their works are read. Which mathematician has studied the works of Euler or Gauss? Which physicist has read Newton's "*Principia*" or Kepler's "*De harmonice mundi*"? Students of history of science, certainly, but practicing researchers, rarely.

Mathematics and science form a rather homogeneous block which grows, or a stream which flows. New discoveries are added, older parts are discarded. With more or less reason, scientists think that the important results of the classics have survived and are implicitly contained in modern textbooks and monographs.

With philosophy it is quite different. Philosophy *is* the history of philosophy. Philosophy has few well–established and uncontested results; it provides questions, problems, approaches and perspectives, which have retained their validity even today. Some problems formulated by Parmenides, Plato, Aristotle, Descartes, or Leibniz have not lost their actuality to the present day. The whole intellectual personality of a philosopher, in a sense, is much more important than the personality of a scientist. Newton's or Einstein's discoveries have merged into the general stream of science. Plato's views are discussed as Plato's views, Hegel's views are commented or condemned as Hegel's views. In his charming little book, Jaspers (1953) says that philosophy is not a science with more or less well established and recognized results, but an ongoing process of thinking: "Philosophy means to be on the road. Its questions are more essential than its answers, and each answer becomes a new question." Sometimes, walking on the road of philosophy, one sees wonderful new vistas, but the true philosopher is not satisfied: he moves on.

Thus all important philosophers from Plato to Wittgenstein are for us, so to speak, *contemporary philosophers*. This may be compared to music: judging from concert programs, available CD–recordings etc., our real contemporary composers are Bach, Mozart, Beethoven, and Brahms! Also the best book on Beethoven is no substitute for listening to his symphonies, and the best book on Plato's philosophy cannot dispense us from reading his dialogues. On the other hand, although Einstein wrote fine books on relativity theory, there are contemporary works such as (Misner et al. 1973) which are more modern and more comprehensive and thus, in a certain sense, do supersede Einstein's books.

Let me conclude with a quotation from C.F. von Weizsäcker (1970, p. 202), which I owe to Viktor Gutmann (Vienna):

> [The physicist does not notice] ... that by rejecting professional philosophy he did not free himself from philosophy but became himself a dilettante philosopher. Unconscious philosophy, however, is in general worse than a conscious one, and thus just the most profound thinkers of modern physics invariably return to philosophy.

Chapter 6

Philosophical implications of science

6.1 Matter and mind

Mein Kind! ich habe es klug gemacht,
Ich habe nie über das Denken gedacht.
Johann Wolfgang von Goethe

There is no doubt that matter and self–conscious mind exist in some way. Different opinions only concern their ontological status (secs. 1.2, 5.1, and 5.2). The clear–cut Cartesian dualism of two independent substances (to a lesser degree the "interactionism" of Popper and Eccles) may suffer from the "fallacy of misplaced concreteness" (Whitehead). An *emergent theory of mind* in the sense of R.W. Sperry (Globus et al. 1976, pp. 163, 181) might be more acceptable to the present habit of scientific thinking, which is perhaps oriented materialistically (physics) rather than idealistically (philosophy). Also dialectic materialism had an emergent theory of mind (cf. Fig. 2.13 on p. 56). The problem with clear–cut materialism and idealism is that modern science has not produced a satisfactory definition, neither of mind nor of matter. We recall Russell's statement that "modern science has made matter less material and mind less mental" and Weizsäcker's saying "Tell me what is your philosophy, and I will tell you how you must define matter".

In fact, the physical definition of matter (mass) as given by classical mechanics (the constant m in eqs. (3.1) and (3.11)), general relativity (eq. (3.55) on p. 96), and especially in quantum mechanics (sec. 3.5) are extremely abstract; see also sec. 5.1, especially Fig. 5.1. In particular, materialist theories of mind that are currently very fashionable suffer from the grave defect that they are based on classical physics. A contemporary theory of brain processes, however, requires a treatment in terms of quantum mechanics, in which the mind of the observer seems to enter in one way or other; see sec. 6.4.

A relatively safe, although crude and perhaps corresponding to current fashion only, analogue for brain–matter and mind is the comparison with computer hardware and software as discussed in various parts of the book. This analogy might even be used as an argument for the "*immortality of the soul*": a computer program may survive even if the computer has become outdated and is no longer used . . . Of course, this is not a *proof* of immortality, only a possible *model*; such "proofs" have been given and refuted in philosophy since Plato's dialogue "*Phaedo*". (Generally, many "proofs" in philosophy and quite a few "proofs" even in science are little more than more or less plausible analogies!)

It is not without interest that St. Thomas Aquinas, the most influential Catholic philosopher, has considered mind the "form" of the human body, in agreement with Aristotelian terminology (sec. 5.4). Thus mind "informs" the body to become a living human being. Cf. expression (4.19 on p. 188) and (Tresmontant 1971). This is far less "substantial" than Descartes' mental substance!

Emotions. It is characteristic that most philosophers concentrate on conscious *thinking*. Emotions and feelings are usually disregarded or considered secondary. This disregard is essential if the brain is to be treated as a computer. The limbic system (sec. 1.1), however, plays a basic role in human consciousness. Feelings must be considered as fundamental as (or even more fundamental than) logical thinking; cf. (Whitehead 1933, Chapter XI) and the end of sec. 2.1.

Selected readings on the body–mind problem may also be found in (Hofstadter and Dennett 1981), (Rosenthal 1991), and (Perry and Bratman 1993).

6.2 Materialism, idealism and the outer world

Materialism and spiritualism,
which are only defined by concepts
taken from each other,
are two aspects of the same thing.
Niels Bohr

It is best to take the reality of the external world as something that is immediately given (Gilson 1972, pp. 286–287), no less directly than Descartes' "Cogito" (sec. 5.4), in agreement with intuitive common sense and with the philosophical tradition up to Descartes. It is only with Descartes and Kant that philosophy has started with the subject rather than the object.

Concerning materialism and idealism it is instructive to compare Hegel with Marx and Engels. If we simply replace Hegel's "spirit" everywhere with "matter", we more or less get dialectic materialism. To a mathematician, what counts is the *logical structure*, and the logical structure is largely the same in all dialectic philosophy.

Now one might object that "matter" *ontologically* is very different from "mind" or "spirit". This may well be, but the mathematical equations which implicitly define matter are supremely indifferent concerning the "nature" of the quantities which these equations relate. For instance, "temperature" is only statistically defined, hence it may well be considered a creation of the human mind. The quantum state function ψ (sec. 3.5) expresses "matter waves", which essentially incorporate our subjective knowledge. So are matter waves objective or subjective? Mathematicians would hardly "mind", for them only the form of the equation "matters" (the reader will kindly forgive me this pun). Let us recall Bohr's saying which served as motto of the present section:

> Materialism and spiritualism, which are only defined by concepts taken from each other, are two aspects of the same thing.

To return to dialectics, it may be said that "matter" must somehow be different from "spirit". However, the "matter" of dialectic materialism has many properties reminding of Hegel's spirit: it is capable of creative evolution, from hydrogen originally present after the "big bang" to intelligent man. This "matter" has all the properties of Spinoza's "*deus sive natura*". So it is largely a question of words. As Findley (1958, p. 23) says: "There is, however, as much materialism in Hegel as in Marx, since matter is for him certainly a *stage* in the 'Idea'." More about this was said in sec. 5.1.

Why, then, so much ado about "idealism" or "materialism"? There is a strong *emotional* difference. "Matter" sounds "scientific", "down–to–earth", and "practical", for the "working class". "Spirit", on the other hand, sounds "aristocratic", "theoretical", and even "religious".

Concerning the reality of the outer world, common sense says that the world exists independently of an observer such as myself. Classical physics and relativity are consistent with this view, but not quantum theory, at least not in the generally accepted Copenhagen interpretation. In fact, quantum theory expresses the interaction of the observer, the measuring apparatus, with the external world rather than describing this world in itself (sec. 3.5).

We can disregard the observer, but then quantum theory gives only statistical averages. If one is satisfied with these statistical averages, in the sense of objective probability or propensity (sec. 3.3), then quantum theory could perhaps be said to describe the external world as such.

If this were not the case, the question would arise *what had happened before man appeared*. Earlier, there was no observer to make the quantum phenomena "definite". Clearly, however, nature took its well–defined course also before man. As Weizsäcker said, "nature was before the appearance of man, but man was before natural science". So the dinosaurs did not have to worry about the course of nature, even without performing quantum experiments. They didn't even need Berkeley's God in order to exist (see the limerick in sec. 1.2 on p. 11).

For further discussion we refer the reader back to sec. 5.1.

6.3 Time, creativity, and the block universe

The creative advance of the world
is the becoming, the perishing, and the
objective immortalities of those things
which jointly constitute stubborn fact.

Alfred North Whitehead

As we have seen in sec. 3.7, there are two different doctrines of : "open", creative and the block universe. The first doctrine says that the passage of is real; at present, the future is not yet determined and can still be influenced; thus there is room for creativity. The second doctrine says that the block universe is essentially finished; everything has been determined beforehand, and there is no place for freedom and creativity; the passage of is only an illusion. According to Weyl (1949, p. 116):

> The objective world simply *is*, it does not *happen*. Only to the gaze of my consciousness, crawling upward along the life line of my body, does a section of this world come to life as a fleeting image in space which continuously changes in time.

Both doctrines have prominent adherents. Curiously enough, the representatives of each doctrine take their view as an absolute and self–evident matter of fact and frequently do not even bother to argue with the other party. Representatives of the block universe are Minkowski (cf. the motto of sec. 3.4), Weyl (just quoted), Gödel (in Schilpp 1949), and Gold (in Duncan and Weston–Smith 1977). The block universe dates back to Parmenides (sec. 5.4).

The majority, however, takes an open, creative universe for granted, also without much discussion: Whitehead, Weizsäcker, and many others. (Einstein seems to have vacillated between the two doctrines.) In fact, if we take entropy, biological evolution, and the freedom of the will seriously, we have little choice. Whitrow (1980) and Popper (1982) thus provide arguments for an open universe.

The first doctrine is perfectly compatible with general relativity (e.g., relativistic evolutionary cosmology), although the block universe would come somewhat more natural to the four–dimensional space– of Minkowski and Einstein, considered as a mathematical manifold.

A synthesis between the two complementary doctrines may have been pointed out already by Boethius (sec. 5.4). For us, the universe is open and ready to be influenced by our actions, in which we are basically free and for which we are responsible. For an observer who lives in the very far future, immediately before the end of the universe and of , the world appears a finished block universe. Only God sees the universe in both complementary ways, according to Boethius.

A detailed physical discussion has been given in sec. 3.7. We remember that the *minus sign* in the metric (3.69) on p. 120, making light cones possible, *is an essential condition for an open universe and thus for creativity and free will* (cf. next section). With a positive–definite metric (3.72), only a block universe would be possible. We also recall that "time travel" is impossible except in fictitious block universes of Gödel type.

Indeterminism and the open universe

If the determinism of classical mechanics (Laplace's demon, sec. 3.1) were correct, then the past would completely determine the future, or even the present would completely determine both past and future. Thus everything is related by a rigorous determinism, there is no place for novelty, and the block universe would be appropriate.

In the same way, an open, creative universe must have some features of indeterminism, so that novelty can arise.

This is particularly evident in the beautiful theory of Whitehead (1933, Chapters XII and XIV), see also sec. 3.8. Here an "actual occasion" (a monad in space–time) has a physical and a mental pole. The past determines the physical pole, but the mental pole provides for creativity and novelty. The action of mind on matter (action of World 3 and World 2 on World 1, sec. 5.2; downward causation, sec. 4.1) is thus entirely natural and immediate.

If one does not accept Whitehead's metaphysics, the one may follow Popper (1982) who provides a comprehensive, profound, and (to me) absolutely convincing argument for indeterminism and an open universe. Popper here anticipates chaos theory: a first draft of his book dates back to 1956. (By the way, Popper refers for instability to fundamental work of Hadamard published in 1898, a reference that is standard in the theory of improperly posed problems (sec. 3.8), but much less quoted in chaos theory!) The following reasoning is based on Popper (1982) and Penrose (1989), as well as on secs. 3.1 to 3.5. (As usual, we do not presuppose knowledge of the references, but they may be helpful later on for a deeper understanding.)

Classical determinism (Laplace's demon, sec. 3.1) presupposes *stability*: small causes produce small effects; a small error in the initial data causes only a small error in the result.

Chaotic systems (sec. 3.2) are unstable: small initial errors may later produce extremely large errors. The classical example is Edward Lorenz' "butterfly effect" in weather prediction. The initial errors may be arbitrarily small (not zero, of course), corresponding to the best accuracy of our measurements. It is clear that any physical measurement is affected by measuring errors: perhaps one part in a billion, but nevertheless never zero.

Already Poincaré (1908) formulated the properties of unstable and chaotic phenomena in an admirable and entirely "modern" way, even including the "butterfly effect":

> If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment. But, even if it were the case that the natural laws had no longer any secret for us, we could still only know the initial situation *approximately*. If that enabled us to predict the succeeding situation *with the same approximation*, that is all we require, and we should say that the phenomenon had been predicted, that it is governed by laws. But it is not always so; it may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have a random phenomenon.
>
> Our second example will be very much like our first, and we will borrow it from meteorology. Why have meteorologists such difficulty in predicting the weather with any certainty? ... The meteorologists see very well ... that a cyclone will be formed somewhere, but exactly where they are not in a position to say; a tenth of a degree more or less at any given point, and the cyclone will burst here and not there, and extend its ravages over districts it would otherwise have spared. If they had been aware of this tenth of a degree, they could have known it beforehand, but the observations were neither sufficiently comprehensive nor sufficiently precise, and that is the reason why it all seems due to the intervention of chance. Here, again, we find the same contrast between a very trifling cause that is inappreciable to the observer, and considerable effects, that are sometimes terrible disasters.

He then goes on to discuss the fundamental importance of these phenomena for probabilistic systems, completely in line with the present book!

Deterministic chaos is a rigorous mathematical consequence of classical dynamics. In a certain sense, "most" classical systems are unstable. This is shown in (Penrose 1989, pp. 174–184) for Hamiltonian systems. Thus, in fact "*classical mechanics cannot actually be true of our world!*" (*ibid.*, p. 183).

As a matter of fact, this statement, shocking as it is at first, cannot surprise us. As we have seen in sec. 2.4, even points in space and their distance (Fig. 2.6 on p. 38) cannot be defined with absolute precision. This is true *a forteriori* for the point masses and similar concepts of classical mechanics. Here we do not even use the fact that classical mechanics is only an approximate limiting case ($v/c \to 0$) of relativity theory and ($h \to 0$) of quantum theory!

Thus the mechanical determinism of Laplace's demon simply cannot be true for our actual world, no matter how much it appeals to our intuition shaped by our first childhood experiences with mechanical toys and by unconscious brainwashing, unnoticed by teacher and student, in physics courses in school and university (I give a university course on classical mechanics myself!).

Mechanics, of course, works (almost) perfectly in the textbook examples of planetary motion, free fall, throwing a stone along a parabolic trajectory, pendulum, etc. These simple examples may be supplemented by equally simple examples *where it does not work*: e.g. flipping a coin or throwing a die as mentioned in sec. 3.3. Nobody would try in practice to predict the outcome of dice–throwing (Fig. 3.4 on p. 84) by a computation of the mechanical trajectory: the outcome is much too unstable.

A reason why we do not spontaneously notice this discrepancy between throwing a stone along a parabolic trajectory (deterministic) and throwing a die (indeterministic) may be that missile trajectories are treated in courses and books on mechanics, and dice–throwing is treated in courses and books on probability, situated in different watertight compartments, at least for many students. (This is why courses on natural philosophy exist: to break down the walls between watertight compartments ...)

In the case of throwing dice, the outcome can only be predicted probabilistically by symmetry considerations: all faces have equal probability p_k.

Here we may still object that even in this case, precise trajectories exist "in principle" (as theoretical possibilities without any practical significance). Very similar practically, but very different theoretically, is the case of *quantum mechanics* (sec. 3.5). Here *only* probabilities p_k (3.60) are meaningful (p. 102); it is meaningless and theoretically wrong to speak of trajectories.

In spite of these theoretical differences, the outcomes of a quantum experiment and of dice –throwing are remarkably similar: only probabilities are meaningful. Einstein expressed his lifelong objection to quantum theory by the famous aphorism: "God does not throw dice" (our motto of sec. 3.3).

6.4 Freedom of the will

The experience of being free
is a real experience.

Euan Squires

The problem of free will has a close relation to the mind–body problem (“downward causation”, cf. sec. 4.1; action of World 2 on World 1, cf. sec. 5.2) and to the concept of an open, creative universe (cf. sec. 6.3).

Immediate perception of free will

In our decisions, actions, and volitions we are influenced by external circumstances, but we feel that in our ultimate decisions we are free. Otherwise there would be no intentional action. Planning would be useless. We would not be morally responsible for our actions. Punishment would be a farce. Policemen and judges could retire immediately, hoping not to be killed, together with other people, by criminals who could act without moral scruples since they are not following their own free will and hence are not responsible for their murders.

A consequent denial of free will, like solipsism (sec. 5.1), is difficult to refute, but still more difficult to maintain in a credible way. By the way, both views have in common that they consider something “normal” as an illusion: free will and the external world, respectively, are regarded as illusions.

Anyway, it seems that people who consistently and honestly deny free will, have been and are being eliminated by biological evolution. Trying to find such persons, the most probable places would be prisons and mental hospitals.

Free will and boundary conditions

Consider Fig. 3.3 on p. 81. Two “trajectories” (possible lines of action) which originally, at time $t = t_0$ are very close, would subsequently follow completely different courses. The difference at $t = t_0$ may be as small as the diameter of an atom; two neighboring trajectories will finally diverge and lead to very different actions.

In this model, free will does not so much affect the deterministic trajectory (any possible action), as the *boundary conditions*, in our case the initial conditions at time $t = t_0$ (Fig. 6.1). The “orthogonality” between the trajectories and the initial line $t = t_0$ is a nice expression of the *complementarity* between freedom and necessity. It is comparable

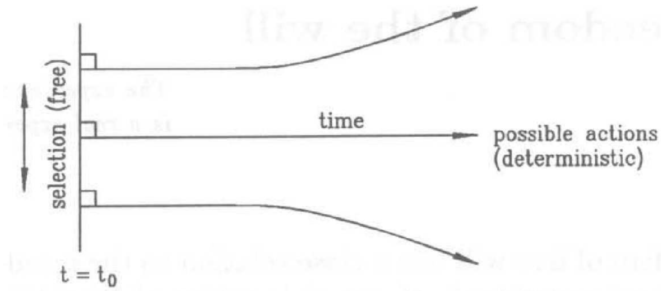


Figure 6.1: Selection (final cause) and action (deterministic cause)

also to the complementarity between "physical laws" (corresponding to action) and "software laws" (corresponding to selection, cf. Fig. 6.1).

Randomness and fluctuation occur everywhere in nature, as we shall see in sec. 6.5. The random fluctuations constantly going on in our neural and brain activity might possibly provide a mechanism for facilitating the free process of selection shown in Fig. 6.1. Random fluctuations will probably save the life of *Buridan's ass*, which, according to (Russell 1945, p. 213) "was unable to choose between two bundles of hay placed at equal distances to right and left, and therefore died of hunger."

Analogy with a thermostat

In sec. 4.1, we have considered a thermostat as a model for "downward causation", the action of mind on the body. There is no need to repeat the argument in detail: the thermostat "causes" the required temperature, mind "causes" the willed bodily movement, software "causes" the computer "hardware" to perform the desired computation. "Causation" here means "downward causation", or Aristotle's "final cause".

These "final causes" generate the differential equations and other "efficient causes" necessary to perform the desired "program" (in the sense of computer language). "Efficient causes" in the sense of Aristotle are simply the deterministic operations of usual causality as embodied, e.g., in the differential equations of classical mechanics. This has been seen to work even in classical physics, following the principles of Fermat and Euler–Maupertuis (ends of secs. 3.1 and 4.1).

The thermostat case of automatic regulation is very instructive also in another respect. In order to get appropriate behavior (left side of Fig. 4.3 on p. 163), we must have "damped" solutions of form (4.1). "Damping", or *friction*, is a typical "dissipative" phenomenon, that is, *mechanical energy is not conserved*. What happens? Part of mechanical energy is "dissipated" by converting it into thermal energy: friction generates heat. So the *total energy*, mechanical plus thermal, *is conserved* after all.

Conservation of energy. It is sometimes asserted that free will is incompatible with conservation of energy. For this reason we have in sec. 4.1 considered the case that the input signal (corresponding to the action of mind) *has no power at all*, giving the example of a compact disc. So the *energy balance* in the brain is not influenced at all. (That the brain uses a great amount of *energy* for its work is no more surprising than the large amount of energy used in a hi–fi amplifier with its loudspeakers or the even larger amount used by a modern supercomputer.)

To be still more concrete, let us consider the following example. If I wish to pick up my pencil, this wish will cause the necessary bodily movements. In the same way, if I put the thermostat of my room to 22 degrees centigrade, this will cause the heating or cooling to act so as to produce the desired temperature. The (very small) energy I need to move the pointer in the thermostat is practically the same whether I select $20°$ or $22°$ (the energy needed afterwards to produce the selected temperature is something different!). In the same way, the (extremely small) energy in the activity of my brain involved in my decision (to pick up or not to do so) is practically the same (the mechanical energy involved in actually moving my hand to pick up the pencil is something different). So thermostats and free wills indeed do not violate the principle of conservation of energy.

Anyway, as all laws, conservation of energy may be only approximate, and it is taken much more narrowly by biologists than by physicists who know about the difficulties of energy conservation in general relativity and quantum theory.

In terms of Fig. 6.1, "selection" corresponds to temperature selection by moving the pointer (almost effortless) and "action" corresponds to the production of the selected temperature (involving considerable energy).

Quantum theory

There is no doubt that quantum theory governs neural activity rather than classical deterministic physics (Stapp 1993). Nonetheless is it controversial whether the indeterminism of quantum jumps is *directly* responsible for free will, as some physicists think; cf. (Jordan

1968, p. 338). The reason is that between quantum jumps, the quantum state function evolves deterministically (according to Schrödinger's equation), and the quantum jumps themselves are governed by statistical laws which seem to be no less rigid than deterministic laws, so as to leave little room for freedom. At any rate, quantum fluctuations provide an omnipresent random background which is important as we have seen above.

Also, quantum theory mitigates the dualism between mind and matter showing that the simplistic materialism of some modern "neurophilosophers" and "neuroscientists" such as Churchland (1988), Dennett, and Edelman (1989) is inadequate: the adequate quantum theory, in some way, always seems to involve mind; see sec. 3.5, (Lockwood 1989), (Margenau 1984), (Squires 1990), and (Stapp 1993).

Going an essential step further, the physicist F. Beck and J.C. Eccles (Eccles 1994, sec. 9) have elaborated a detailed quantum–mechanical model for mind–brain interaction. Such a model may not describe the real situation, but it shows *that such an interaction is possible*.

This model is based on the plausible assumption that mind may change the distribution of quantum probabilities of micro–events in the brain, whose combined action produces the desired macro–event. This is another example of *amplification* (sec. 4.1).

The argument of Lucas

In sec. 5.2 we have outlined an ingenious argument of Lucas (1970, §25). He uses *Gödel's theorem* as an additional argument that human thinking is not determinable by any physical law: the human brain cannot be a deterministic machine or computer because human thinking essentially transcends computability (sec. 2.3). As for almost all philosophical arguments, there are also counterarguments (*ibid.*, §26) which can again be refuted (§27).

The argument of Epicurus

Following Lucas (1970, §21) we can also apply a variant of the Epicurus–Haldane argument already given in sec. 1.2. *If everything is determined, then it is also determined that I believe that everything is determined.* I thus believe this not because I find *by logical considerations* that it is true, but because the structure of my mind has been determined *by physical law* so that I *must* think this way (regardless of what logic says). Many things, however, are determined by physical law which need not be true. Hawking (1993, p. 129) says:

Each week, my mail contains a number of theories that people have sent me. They are all different, and most are mutually inconsistent. Yet presumably the grand unified theory has determined that the authors think they were correct. So why should anything I say have any greater validity? Aren't I equally determined by the grand unified theory?

What does Epicurus' argument really show? It does not directly refute determinism. It shows, however, that *universal determinism* ("everything is predetermined by a grand unified theory" (sec. 6.6)) *is non–arguable*. If someone claims that he has found a proof of universal determinism, we can immediately reply that his reasoning is inevitable if universal determinism is true, regardless of the argument being logically valid or not. So his argument is simply worthless, as the above quotation from Hawking shows very clearly.

The argument of Epicurus is particularly malicious and insidious. It turns determinism against itself: it beats determinism at its own game.

By the way, Gödel's proof uses *self–reference* (reflexive logic, see end of sec. 2.5), and also the argument of Epicurus is self–referential. Otherwise, both arguments have little in common, but we see again the importance of self–reference in philosophic reasoning.

Indeterminism is not enough

As Popper (1982, pp. 126–127) has pointed out, "indeterminism is not enough" for freedom of the will, "indeterminism is *necessary but insufficient*". What does this mean?

As we have seen in sec. 6.3 ("indeterminism and the open universe"), indeterminism rather than determinism seems to reign in the universe: we have the instability of chaotic systems, not to speak of quantum indeterminacy. The cause does not determine the result to any useful accuracy; unmeasurably small variations (errors) in the initial data may completely change the result: the "butterfly effect", cf. Fig. 3.3 on p. 81.

Thus there is practical indeterminism, but how can we "*harness*" (Penrose 1989, pp. 172, 174) this indeterminacy to get free will? To be sure, free will is incompatible with rigorous determinism, but indeterminism does not yet directly imply the freedom of the will: *indeterminacy is not enough*. Does the uncertainty of a weather forecast define, or at least increase, my freedom in planning a weekend hike in the mountains? This is an obvious absurdity.

What indeterminacy does, however, is *to create a free space in which the will can act*. Indeterminism is necessary but insufficient: it provides the necessary leeway, but only the action of mind is then sufficient for an act of free will. Mind belongs to World 2, partly even to World 3 (sec. 5.2), but our actions refer to the material World 1. For free will it is therefore necessary that *World 1 is open to World 2* (Popper 1982, p. 113 ff.). This action of mind on matter has been called *downward causation* (see above).

Let us give an example: unfortunately, tomorrow I have to attend a faculty meeting. The meeting has been scheduled long ago, it is known to all professors, and it has even been announced in the University Bulletin. Being objective or at least intersubjective, the scheduling of the meeting belongs to *World 3* (although Plato may not like to see it in his lofty realm of ideas). I see the schedule and reluctantly make up my mind to attend the meeting. This decision clearly belongs to *World 2*: World 3 thus has acted on World 2! The next day, I really attend the meeting and walk to the meeting room. This movement of walking obviously belongs to *World 1*. Thus we have the action: World 3 (objective schedule) → World 2 (subjective decision) → World 1 (physical movement).

How can this interaction between World 2 and World 1 be realized? We consider three possibilities, summarizing what we have already said.

(1) *Whitehead's solution.* In Whitehead's metaphysics (Whitehead 1933, Chapter XII), see secs. 3.8 and 6.3, an actual occasion (a monad in space–time) has a physical pole and a mental pole, which are more or less on the same footing. Hence elements of World 1 and World 2 are integrated in a most natural and spontaneous way.

(2) *Thermostat analogy.* As we have just seen, a thermostat causes the room temperature to assume a prescribed value. The temperature selection (cf. Fig. 6.1 on p. 246) is made by moving a pointer on the thermostat, which can be done with arbitrarily small energy (by appropriate lubrication, amplification, etc.). The corresponding room temperature change involves very considerable energy (heating or cooling).

It would thus seem that we have the case of a small cause resulting in a large effect, which, according to chaos theory, corresponds to instability and uncontrollability. This is certainly not true in the present case which is very precisely controllable!

The explanation is *feedback*. The temperature achieved is constantly compared to the desired temperature (fed back to the thermostat), and the temperature difference ("deviation" or "error") is used to improve the room temperature until the error is ideally zero.

Thus feedback allows "controlled amplification", in contrast to the "uncontrolled amplification" in unstable chaotic systems. A particularly impressive case is a hi–fi equipment amplifying the almost microscopic signal on a compact disc to produce the majestic sound of a Beethoven symphony; again feedback is, of course, essential. All *servomechanisms* operate in this way, as we have seen in sec. 4.1.

A final example may illustrate this. I wish to walk from point A to the nearest village. From point A I see the village and select the direction accordingly. In the first case, there is a dense forest between A and the village, so the best I can do is to maintain the initial direction as precisely as possible. Not surprisingly, I will find out that I have missed the village by a couple of kilometers. In the second case, there is no forest and I can see the village all the time, so that I can continuously correct my direction so as to arrive at the desired place. The first case corresponds to *instability* (a small initial error is enormously magnified), in the second case we have *feedback* (by seeing the village) continuously stabilizing the direction of my walk.

Is there a more direct relation between automatic regulation (thermostat, servomechanism), and indeterministic, chaotic motion? As we have seen above, feedback must involve *damping* or *friction*. Friction, however, converts mechanical energy into *heat*, which is nothing else than the chaotic motion of gas or liquid molecules as described by statistical mechanics (sec. 3.2)! In sec. 4.1 we have seen that damping is inseparably related to positive feedback which always occurs in automatic regulation.

Thus, in Penrose's terminology introduced above, automatic regulation "*harnesses*" chaotic motion. This equally holds for "indeterminist" classical chaotic motion and for the "indeterminism" of quantum mechanics. (Heat in solid bodies corresponds to vibration of molecules, which is a quantum rather than a classical effect!)

(3) *Quantum interaction.* Above we have mentioned the model of Beck and Eccles (Eccles 1994). The change of probability distributions may be regarded as an analogue to moving the pointer of a thermostat.

6.5 Laws of nature

God has not made the original equations available.

James Gleick

Take the example of Kepler's three laws for the motion of planets around the Sun. The first and most important law is that planets move along ellipses whose focus is the center of the Sun. It is easy (I do it regularly in my introductory course on mechanics for students in the second semester) to derive, from Kepler's laws, Newton's laws of gravitation.

Inversely, Newton's law of gravitation permits the derivation of Kepler's laws, with two important qualifications.

(1) Possible orbits are not only ellipses, but also other conic sections: parabolas and hyperbolas. This is relevant for comets.

(2) If, in addition to the Sun and the Earth (say), the attraction of other planets is taken into account, then the orbits are slightly perturbed: *they no longer are exact ellipses*.

The second fact is particularly important: Kepler's laws do not hold exactly, due to the perturbation of other planets! If the measurements of Tyho de Brahe, from which Kepler derived his laws, had been more accurate, then Kepler would never have arrived at his simple laws, and Newton's laws would not have followed so readily. In this case, as Alfred North Whitehead has pointed out, too accurate measurements might have hindered the progress of science!

This is another nice instance of a Hegelian triad (sec. 2.5). Thesis: Kepler's laws; antithesis: Newton's laws; synthesis = thesis on a higher level: Kepler's laws with corrections for the effect of other planets.

Even Newton's law of gravitation, however precise (to 10^{-7}), is not absolutely exact: Einstein's theory of general relativity is better (sec. 3.4). Einstein's theory comprises Newtonian mechanics as a special case for "small" velocities and "weak" gravitational fields, which is sufficient for most applications.

Einstein's general theory of relativity is extremely elegant, general, and beautiful. So far, it has never been refuted by experiment. Is it *absolutely* true?

Quantum mechanics (sec. 3.5) is equally elegant and general; it has been confirmed by experiments better than any other theory. Is it *absolutely* true?

Quantum theory holds for very small distances: between nucleus and electrons in an atom, and between atoms in a molecule. At these small distances, quantum effects *must* occur: all physicists agree on that.

Now, no quantum effects can be derived from general relativity. Therefore, general relativity cannot hold for very small distances. Thus it cannot be *absolutely* true.

But many scientists agree that general relativity is perfect and unique with respect to elegance and conceptual simplicity.

How can it happen that such a beautiful theory is not *exactly* valid? Let me try to illustrate this by means of examples which we understand better. The ancient Greeks thought that the Earth is a perfect sphere because the sphere is the most perfect and symmetric of all surfaces. In fact, a fluid planet without rotation must be an exact sphere. A *rotating* planet, however, is slightly flattened and very similar to an ellipsoid of revolution. But the actual figure of the Earth, the geoid, is not an exact ellipsoid: it deviates from an ellipsoid by 100 m at most and is rather irregular, because of the attraction of the irregular "topographic masses": mountains, hills, valleys, every feature which makes our planet attractive and interesting. So symmetry (sphere, ellipsoid) creates the basis, but irregular features are superimposed.

In the case of planetary motions, Kepler was so fascinated by the beauty and simplicity of his laws, especially of the fact that the planetary orbits were exact ellipses, that he never ceased to praise the divine harmony of the Universe.

Nevertheless, Newton introduced perturbations to this perfect harmony, and Einstein replaced Newtonian mechanics by an even more accurate, general, and beautiful theory.

What should we conclude? Elegance and internal perfection may be a necessary condition for a good scientific theory; they are, however, not sufficient. (Wolfgang Pauli exaggerated when he said: "Elegance is only for the tailor", implying "not for the scientist".)

The very perfection of both general relativity and quantum theory may stand in the way of their unification, but many physicists believe that such a unification is necessary and also possible. In his extraordinary book, Penrose (1989) has attempted to fathom the important and far–reaching consequences which such a unification would have even for general human thinking and for artificial intelligence. (*Special* relativity and quantum theory are, of course, unified in Dirac's theory of the electron.)

Let us come back to our question: *What is a physical law?* Does it reflect the behavior of nature or only the structure of our mind and of our perceptional apparatus? This question was considered already in sec. 1.4 of our book, especially in connection with Eddington's example. The answer is: probably both, it has an objective and a subjective component. The subjective component is particularly evident in quantum theory (sec. 3.5), but no physicist really doubts that his laws do have an objective basis in nature. What would happen if the engineer calculated bridges or high buildings according to formulas which do not express real properties of nature?

Are natural laws exact?

Kant has considered Euclidean geometry and Newtonian mechanics as a priori given and hence exact. We now know that even symbolic logic (set theory) and mathematics are under suspicion (logical paradoxes, Gödel's theorem, see sec. 2.3). The objects of everyday life, including physics, are subject to "fuzzy logic" (sec. 2.4). In the same section, we have seen that measuring errors are unavoidable, and attempts to idealize one concept (regarding the angles α and β as errorless) may lead to a falsification of others (the angle γ). Furthermore, we have just seen that even such apparently "perfect" laws as Kepler's laws or Einstein's theory of general relativity cannot be valid absolutely.

The beautiful and important laws of thermodynamics (including the famous Second Law of Thermodynamics) are a consequence of applying mechanics to a system of a huge number n of particles or molecules (statistical mechanics) and, strictly speaking, require $n \to \infty$.

Properties such as *temperature* are averages over a large number of particles (it would be meaningless to speak about the temperature of a system that consists of 10 molecules only). At least thermodynamics is only a *statistical* theory, an example of "order out of chaos".

Quantum measurements are limited by *Heisenberg's uncertainty relation* (sec. 3.5), and complex mathematical and logical theories might be affected by the limitations imposed by *Gödel's theorem*.

Hence we find inaccuracies, errors, and limitations wherever we look. It becomes increasingly more difficult to believe that physical laws are exact if already the concepts which they relate, cannot be defined with absolute accuracy.

Randomness and fluctuations occur over and over again in nature. On the quantum level we have a universal background of essentially random *quantum fluctuations* (see end of sec. 5.2). But also on the classical level we have *Brownian motion*, the erratic movement ("random walk") of a small dust particle suspended in a fluid and of course, the *statistical mechanics* of air molecules which we have just mentioned.

Even in our nervous system and in the brain we seem to have some "randomizers", similar to "random number generators" familiar from computer work. For this reason, repeated precise measurements never give the same result: there are *random errors* described by the Gaussian curve of Fig. 4.12 on p. 184. Or try to follow a precisely drawn circle by free tracing performed by your hand: you will never be able to stay exactly on the curve (Fig. 6.2). No matter how firm your hand is, it will inevitably perform (very small) shaky movements. There seems to be a randomizer in our nervous system which, however, works on the physical level only and does not affect the mental level (sec. 5.2).

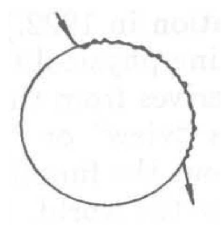


Figure 6.2: Tracing a circle with a shaky hand

The great contemporary physicist John Archibald Wheeler even questions the unlimited validity of physical laws in space and time. Are the laws of physics on Sirius the same as on Earth, and have the laws governing the Big Bang been the same which we know today? If they are (and so far we have hardly a reason to doubt this although it appears to be short of a miracle to some people) we could never test this by direct observation. In a certain sense, Wheeler (1994, p. 300) even says: "No laws." Also: "Law without law."

What has been left of "the eternal laws of nature to which it is inexorably subject"? I don't know. In the absence of man, are there still laws of nature? Obviously. The same? Probably not. Why are "our" laws so applicable and beautiful? What about Wigner's "unreasonable effectiveness of mathematics in natural sciences" (sec. 3.9)? Why cannot we then have exact laws? Why "did not God make available the exact equations"? Are there "exact equations"? Questions, more or less legitimate ones, but no answers ...

Generally, because of unavoidable measuring errors, it is never possible to verify or falsify our physical theories *exactly*! This seems to have been frequently overlooked in discussions on empirical verification or falsification of physical theories (cf. sec. 3.9).

As an example, take the empirical verification of Euclidean geometry by measuring the three angles of a large triangle formed by light rays such as done in geodesy. This idea was already proposed by C.F. Gauss. Let us assume, for a moment, that we can define the points ABC of a triangle and measure the angles $\alpha\,\beta\,\gamma$ *with absolute precision.* Will their sum be π or 180°?

Surprisingly enough, the answer is *no*! The reason is that general relativity, which we have no motive to doubt in this context, states that light rays are not straight lines but are slightly curved because of the curvature of space–time due to gravitation. In this sense, empirical geometry is not Euclidean. (This has been hard to accept for many scientists such as Alfred North Whitehead who developed an alternative relativistic theory of gravitation in 1922.)

An original idea concerning physical theories is due to Bohm (1980, p. 3). The word "theory" derives from the Greek "*theoria*" (same root as "theater"!) which means "view" or "perspective". Thus theory is not so much a *knowledge* about the functioning of the world, but rather an *insight*, a way of looking at the world. It is thus primarily an activity of the human mind and as such *a priori* in the terminology of Kant. It is not, however, absolute and errorless as Kant thought, but subject to our errors and weaknesses. One theory, one perspective, may be better than another (relativity is better than classical mechanics), but no theory can be expected to be absolute and final: we cannot expect to catch the whole universe in one perspective. Thus a theory is to the mind what a microscope or telescope is to the eye. Regarding nature equipped with a theory is similar to looking at a landscape through the eyes of one's favorite painter.

Also Whitehead (1933, Chapter VII) considers the possibility that natural laws only hold approximately (somewhat as the criminal law governs human societies). This comparison immediately raises the question: If the laws are only "approximately respected", who sets the limits of possible deviation from the law and watches them? God as the "policeman of nature"?

Conventionalism

Well, if the relativistic concepts differ so little from the corresponding Euclidean concepts, why not assume Euclidean geometry to be exactly valid, and treat the very small deviations (of order 10^{-8}) as "relativistic corrections"? In this spirit, why should we not say that light rays are not Euclidean straight lines but are slightly curved, so that $\alpha + \beta + \gamma = 180° + \epsilon$, where ϵ is a small quantity which is not zero because of curvature effects?

This is the point of view of *conventionalism*. It says that the choice of geometric and physical laws, to a certain extent, is arbitrary or conventional. We know that modern physics considers the speed of light c, given by (4.13) on p. 184, as a fixed errorless constant by defining length in terms of c, cf. (4.17), that is by *convention*. So why not keep Euclidean geometry and Newtonian mechanics and treat deviations as corrections?

Logically, this is perfectly possible, and even practically one frequently proceeds in this way, cf. (Moritz and Hofmann–Wellenhof 1993, p. 255). The great mathematician and author of charming philosophical books, Poincaré (1902, 1908, 1958, 1963) thought that one would never give up Euclidean geometry because of its unrivalled simplicity!

Even great scientists such as Poincaré and Whitehead can be wrong. In fact, Einstein's theory of relativity is such a marvellously perfect theory that one has even preferred to give up Euclidean geometry in order to keep it.

As a matter of fact, the general theory of relativity, in view of its very perfection, has challenged many scientists to invent gravitational theories of similar perfection and similar beauty. Some of them have been refuted by experiment, but others have withstood all tests. Nevertheless Einstein's theory is preferred almost unanimously: because of its unsurpassed perfection, beauty and internal consistency. Remember Wigner and Penrose!

Hence conventionalism is right in the sense that for many theories it is possible to invent rival theories which are equally compatible with experiment, within the limits of measuring accuracy. Here Kuhn (sec. 3.10) holds out against Popper (sec. 3.9): falsification of a theory is only possible within the limits of measurements; new "paradigms" are introduced by other reasons (elegance, utility) in a more or less conventional way.

Positivism

Hipparchus (around 146–127 B.C.) and Ptolemy (2nd century A.D.) explained the motion of the planets as seen from the Earth, by a superposition of *epicycles*. Mathematically this is a series of cosines of form

$$x = a_1 \cos(\omega_1 t + \beta_1) + a_2 \cos(\omega_2 t + \beta_2) + a_3 \cos(\omega_3 t + \beta_3) + \cdots \quad (6.1)$$

Here a_i, ω_i and β_i are constants, and the geometrical explanation is shown in Fig. 6.3.

Thus it corresponds to a superposition of circles C_i of radii a_1, a_2, a_3, ... The angular velocity of the movement of the center of C_2 around the basic circle C_1 is ω_1, the angular velocity of the movement of the center of C_3 around C_2 is ω_2, etc. Finally, the β_i are appropriate phase angles ($i = 1, 2, 3, \ldots$). This is called a system of epicycles. It gives merely a geometrical illustration and interpretation of eq. (6.1) and does not correspond to the actual movement of the planets.

Nevertheless, these theories of epicycles described planetary motion with respect to the Earth, with astonishing accuracy, the better, the more terms (circles of radius a_i) were superimposed (also $i = 4, 5, 6$, etc.)

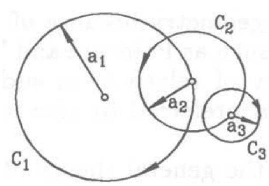


Figure 6.3: A system of epicycles

Only Nicolaus Copernicus (1473–1543) placed the Sun, rather than the Earth, at the center of the universe, and Johannes Kepler (1571–1630) recognized that the planetary orbits are ellipses rather than circles. Finally, as we have seen at the beginning of this section, Newton derived from Kepler's laws his law of universal gravitation, eq. (3.11) on p. 74.

It is now highly remarkable that if the laws of planetary motion, under the influence of the Sun and of the other planets, are calculated from Newton's law (3.11), *referred to the center of the Earth*, the solution has exactly the form of eq. (6.1). Computationally, Newton brings us back to Ptolemy! (By the way, from the convergence or divergence of astronomical series of form (6.1), Poincaré discovered chaos theory, sec. 3.2 ...)

From a *positivist* point of view, the purpose of a physical law is to *describe* (as accurately as possible) rather than to explain. To *explain* using entities that are not directly observable is considered more or less *metaphysical* (see end of sec. 1.2) and should thus be rejected.

Now the epicycle theory, expressed by eq. (6.1), furnishes a perfect and rather simple mathematical description, and the planets have been inaccessible to man until now, as the Moon was before the astronaut Armstrong stepped on it in 1969. It can be argued that accessibility can also be achieved through instruments such as telescopes, so that the planets are directly observable entities, but these arguments need not be accepted by a stubborn opponent. Thus I should say that those who regard the planets as metaphysical entities are still less in number than those who regard in this way atoms, electrons or quarks (have you ever seen a quark?), but the principle persists (have you ever been on Mercury?).

In reality, of course, Kepler's laws of planetary motion, and Newton's and Einstein's (sec. 3.4) theories of gravitation are universally preferred by physicists. This shows that even physics contains *unobservable quantities* so that the positivist point of view is not carried through completely in scientific practice.

Whitehead (1933, Chapter VIII) has given a beautiful example. Earlier in this century it was remarked that a series (6.1) as known before had to be supplemented by extremely small additional terms in order to achieve best agreement with very precise contemporary observation. "Every Positivist must have been completely satisfied. A simple description had been evolved which fitted the observed facts" (*ibid.*, p. 127). But, in fact, the American astronomer Percy Lowell was not satisfied. He found out by computation that the new anomalous terms corresponded to the attraction of an imaginary small planet, an astronomer at the Lowell Oberservatory photographed in 1930 the sky in this direction and found ... the new planet Pluto! Whitehead wrote

> The civilized world has been interested at the thought of the newly discovered planet, solitary and remote, for endless ages circling the sun and adding its faint influence to the tide of affairs ... The speculative extensions of laws, baseless in the Positivist theory, are the obvious issue of speculative metaphysical trust in the material permanences, such as telescopes, observatories, mountains, planets, ...

Thus positivism, with its justified emphasis on observability and mathematical description, is important but should not be interpreted too narrowly.

Here positivism, similar as conventionalism before, represents a certain general direction of thinking rather than a definite school of philosophy such as logical positivism (sec. 5.4). Whitehead's pointed remarks should be understood in this sense.

Laws, theories, and mathematical models

The borders between these concepts are ill–defined and shifting, so that only rough, "fuzzy" definitions are possible.

A *theory* usually is comprehensive and well confirmed by experiment, e.g., Newtonian mechanics, special and general relativity, quantum theory, or biological evolution.

A *law* expresses an important and generally valid feature, formula, or theorem of a theory, e.g. Newton's law of universal gravitation, or the law of conservation of energy. Often, "laws" is practically the same as "theory" or "theories", e.g., when we speak of the "laws of physics".

A *hypothesis* is theoretically possible and interesting, but usually not (yet) completely confirmed by experiment. One frequently speaks of a *working hypothesis*. I should consider the reality of the external world an example of an extremely important working hypothesis.

A *mathematical model* usually implies a simplification to make a phenomenon accessible to mathematical treatment. Examples are the sphere or a slightly flattened ellipsoid of revolution as models for the shape of the Earth which, of course, is more complicated. A perfect example of a mathematical model are Kepler's laws for planetary motion along ellipses which are mathematically very simple and useful but not entirely precise as we have remarked at the beginning of this section.

Very important is the concept of *reference model*, which serves as a basis for determining or defining a real phenomenon, thus

$$\text{phenomenon} = \text{reference model} + \text{deviations}\ .$$

For instance, an Earth ellipsoid serves as a reference for determining the actual Earth surface (the geoid). Kepler ellipses serve as references for determining deviations of planetary or satellite orbits from such ellipses. (The Traffic Code is a reference model for actual traffic (in order to be able to define and punish deviations), but this, of course, is beyond mathematics and even science.)

Mathematical models reach from well–established, purposely simplified mathematical laws to mathematical working hypotheses. Examples for the first type are Kepler's laws and e.g., the ideal gas laws in thermodynamics; examples for the second type are the quantum mechanical model of mind–brain interaction of Beck and Eccles (Eccles 1994, Chapter 9) and the overdetermined and underdetermined linear mathematical models at the end of sec. 3.8.

Penrose's classification. Penrose (1989, p. 152) gave a very useful classification of physical theories. Such a theory may be

SUPERB,
USEFUL, or
TENTATIVE.

SUPERB are, for instance, classical mechanics (within the usual limits of applicability defined by the theories mentioned next), special and general relativity, and quantum mechanics. USEFUL are, e.g., the gauge theories in elementary–particle physics, in particular the "Standard Model" (sec. 3.6). TENTATIVE is, e.g., superstring theory as a possible candidate for a "Theory of Everything" (sec. 6.6).

What is a law of nature?

What do we mean by a law of nature? We have asked this question and found many partial answers in sec. 3.9 and elsewhere, and in particular, of course, in the present section. We keep asking the same questions over and over again!

Never mind, let us keep insisting: if the poet speaks of the "eternal laws of nature", if the reductionist wishes to reduce biology to the "laws of physics", what are these laws? All "professional" answers we have found sound unconvincing. A theory confirmed by induction or surviving falsification is a human construct rather than a law *of nature*, not even *of physics*: it is a law *of man* containing, as we hope, some objective elements. According to Kuhn, it is a currently accepted *paradigm*, some kind of "scientific fashion".

As we have said above, a law of nature, in the ideal case, has an objective and a subjective component, and it is only approximate.

Probing obstinately further, we do not wish the subjective component. We wish a law of nature to be a law *of nature*. Einstein objected to quantum mechanics for this very reason.

Let us try to view the question from another angle. Can we suppose that nature *has its own laws*? I think all scientists believe this, at least subconsciously. Without this belief there would be no physics, no induction, no verification, no falsification, not even a reasonable paradigm. To state it bluntly, we believe that nature, or God (Spinoza: *deus sive natura*), does have its own laws, which are essentially the same as they were before the appearance of man. In terms of James Gleick's motto at the beginning of the present section, we believe that God does have the original equations, even if He does not provide them to us. Nature is ruled by its own laws, and we try to find them, at least approximately. Our laws of nature are approximations to God's laws of nature, and our laws of physics are approximations to God's laws of physics. If you do not like this very "unreductionist" terminology, replace "God" by "nature".

Is this belief justified? Clearly, it cannot be proved. Quantum theory even seems to flatly contradict it. In sec. 3.5 we have seen, however, that “objective” interpretations of quantum mechanics may not be impossible. Following Kant, many philosophers (Neo–Kantians, perhaps the late Wittgenstein) and scientists (Wheeler, Maturana, Varela) seem to question the existence of an external world independent of the observer.

So can we believe that purely “objective” laws of nature exist? I think so. The main argument: We are part of nature, and our thinking should be at least in basic agreement with nature (otherwise we might have been eliminated by evolution long ago). More philosophically, Spinoza says: “*Ordo et connexio rerum idem est ac ordo et connexio idearum*”, the order and connection of things are the same as the order and connection of ideas. Again, we recall the quotation from Penrose at the end of sec. 3.9.

Many things in philosophy, in daily life and in science must be believed. So why not believe in the existence of an external world governed by laws? This belief cannot be proved, but it is not irrational: it has never been contradicted by experience. Science, rather than disproving our naïve faith, has refined it: the external world may be essentially (following Kant, for us) three–dimensional, but the corrections due to quantum theory (possible subjective elements, nonlocality, etc.) and perhaps to some higher–dimensional “super–theory” should be accepted.

The complexity of the concept of physical law

It may be appropriate finally to compile some problems concerning laws of nature, in particular of physics, which we have noticed so far.

(1) How is a law found?

- induction, verification, falsification (sec. 3.9)
- the “unreasonable efficiency of mathematics”: the role of mathematical simplicity, symmetry, and elegance (secs. 3.6, 3.9)
- the role of conventional paradigms (sec. 3.10)

(2) The character of a law

- a priori vs. a posteriori (secs. 1.4, 5.3, 5.4)
- objective and subjective elements: objective truths vs. subjective perspectives (secs. 3.5, 5.3)
- positivism or metaphysics: description vs. explanation

– Do the laws remain unchanged throughout space and time?
– Are laws exact? Do they contain probabilistic elements or uncertainties of Gödel or Heisenberg type?
– Are our laws at least approximations to unknown (and possibly unknowable) "exact laws"?

These and many other questions arise. Some may be answered, some not. At any rate, there does not seem to exist a unique answer to all these questions. A "law of nature" is a "chameleon concept" (J.R. Lucas), it subtly changes its meaning during the course of a discussion, and even more so during the course of a book like the present one. So to speak, this is a "skeleton in the cupboard" hidden in almost any book on science or philosophy of science, and it is best to open the cupboard and study the skeleton.

Let reductionist biologists look for simple physical laws underlying biological complexity: philosophers of science will continue to be concerned about the conceptual complexity underlying the very meaning of a "simple" physical law. Both parties treat meaningful problems in meaningful ways.

A modern treatment of laws of nature, simplicity and complexity, reductionism and emergence is found in (Cohen and Stewart 1994). Also (Barrow 1988) can be recommended. Writings of great scientists about laws of nature are always fascinating: (Eddington 1939, 1959), (Einstein 1957), (Heisenberg 1955, 1958, 1973), and many others quoted elsewhere in this book.

6.6 Theories of everything

Dear Reader, you have before you a Theory of Natural Philosophy deduced from a single law of Forces.

Rudjer Bošković

Your theory is crazy, but probably not crazy enough to be true.

Niels Bohr

An old dream of theoretical physicists, essentially the dream of "Laplace's demon" (sec. 3.1), is to find the "world formula", combining relativity and quantum theory and unifying the four basic forces of physics:

gravitation,
electromagnetism,
weak force,
strong force.

Gravitation and electromagnetism are more or less familiar. For the rest, let us briefly recapitulate from the rather technical Section 3.6. A *free* neutron disintegrates within several minutes, leaving a proton, an electron, and a neutrino. For such and many other transmutations of particles ("radioactive decay"), the *weak force* is responsible. An atomic nucleus consists of protons and neutrons. The electromagnetic repelling force between the positively charged protons must be overcome in order to keep the nucleus together. This is done by the *strong force*. The strong force similarly acts between the quarks that make up protons and neutrons.

Already Einstein had dreamed of the unification of gravitation and electromagnetism (the only forces known at that time) by modifying the equations of general relativity. A very interesting theory was found by Hermann Weyl as early as 1918; it served as the ancestor of modern gauge theories.

Almost all famous theoretical physicists tried their hand at finding unified theories. Einstein worked for most of his life on this problem, unfortunately in vain, but also Eddington in 1923 and Schrödinger in 1950 were fascinated by this question.

Whereas these scientists based their efforts on general relativity, Heisenberg in 1959 tried to proceed from quantum theory. Also his attempt failed.

In the meantime, C.N. Yang and R.L. Mills had formulated the first *gauge theory* in 1954, generalizing Weyl's ideas. Since then, gauge theories have become basic mathematical tools for describing elementary particles.

The first unified gauge theory for electromagnetic and weak forces by A. Salam, S. Weinberg, and S. Glashow brought its authors the Nobel Prize in 1979. Contemporary unified gauge theories include, in addition, also the strong force. Thus we have the *Standard Model* as a preliminary but extremely useful approximation to an ideal GUT, or *Grand Unified Theory*.

Thus three of the four forces have been unified. The great exception is gravitation, which somehow refuses being included in an ordinary gauge theory. An extension of the latter, *supersymmetry*, may provide a possibility to include gravitation as well. A final refinement (and complication) is achieved by regarding elementary particles not as pointlike but as an extremely small ($\sim 10^{-13}$ cm) curve, or loop, a "string".

These strings can be combined with supersymmetry to give *superstring theory*. This theory (or its variants) is so complicated that the mathematical consequences have not yet been elaborated. Thus it has not yet been possible to really understand whether and how well this theory applies to the real world. With superstring theory we are now (1994) so far as with Heisenberg's unified theory in the sixties.

Excellent Rabbi type 1 references on these fascinating topics are (Barrow 1991), (Davies 1984), (Davies and Brown 1988), (Hawking 1988, 1993), and (Kaku 1994). We again refer also to sec. 3.6.

Rudjer Bošković (1711–1787). "One of the most remarkable and neglected figures in the history of modern European science was Roger Boscovich" (Barrow 1991, p. 17). A Jesuit from Dubrovnik, he worked in philosophy, mathematics, physics, geodesy, and similar sciences which at that time were not yet so separated and specialized as today. Recognized by Niels Bohr and Werner Heisenberg, he nevertheless is little known outside Croatia. He proposed a grand unified force law, generalizing Newton's law of gravitation to include all other physical forces. He believed in atoms and elementary particles and tried to find a single law encompassing gravitation as well as the forces that would hold his hypothetical elementary particles together. "His continuous force law was the first scientific Theory of Everything" (Barrow 1991, p. 18).

Bošković is important not so much for his particular contributions to science, but he was a visionary whose ideas influenced great physicists such as Faraday, Maxwell, Boltzmann, and Lord Kelvin. Some of his philosophical ideas are relevant even today, cf. sec. 4.4 (p. 186) and sec. 5.2 (p. 210).

Philosophical problems. A future TOE (*Theory Of Everything*) poses very difficult and sometimes contradictory problems which are, however, profound and fascinating.

(1) In the ideal case, it would describe and predict all physical processes, however complex and "chaotic": laws together with boundary conditions.

(2) If chemistry and biology are, in fact, nothing but theories of particularly complex physical systems, a TOE must also fully explain chemistry and biology, including "self–organization".

(3) It must incorporate the uncertainty principles of Gödel and Heisenberg.

(4) If mental brain activities were, in reality, only physico–chemical processes, then TOE would also describe human thinking; the "freedom of the will" would then be nothing but an illusion.

(5) The TOE, predicting *everything*, must also derive *itself*.

It is clear that all this appears quite improbable. We would then have Laplace's demon (sec. 3.1) again, in contradiction to chaos theory and Heisenberg's uncertainty relation.

As we have seen in sec. 4.5, a reduction of biology to physics may work in this "analytic" direction. In the opposite ("synthetic") direction it probably will not be possible because *information* is needed, in agreement with the "equation"

$$\text{life} = \text{matter} + \text{information} .$$

As a TOE, at least as it appears now, does not contain this "vital" information, "synthesis" and prediction appear more or less impossible.

So the *freedom of the will* does not seem to be in danger, not even from a TOE. Even if theoretically, provided we knew the required initial conditions fully and with absolute precision, we could calculate human thinking, this cannot be done practically because the amount of calculation exceeds all that is humanly possible. Furthermore, if we could exactly predict our future decisions, they would be known, and *this knowledge might change our future thinking and acting*! This ingenious argument was given by Hawking (1993, p. 135). In fact, this "pre–cognition" would constitute an inadmissible *time loop* (sec. 3.7, Fig. 3.20 on p. 129).

Finally, even if a TOE were able to derive itself, how would we know that it is *true*? Of course, in this case, the TOE would be *internally* consistent, but how do we know that it corresponds to *external* reality? Hawking's answer is Darwin's theory of evolution: animals behaving not in agreement with reality, and early man having wrong "theories" about reality are unfit to survive and will be eliminated by evolution (cf. sec. 1.4). Interesting as this argument is (we have used it frequently), I should not consider it a philosophically fully satisfactory answer.

This is related to the argument of Epicurus (sec. 6.4); and also the argument of Lucas (*ibid.*) is of direct relevance here.

Let us therefore ask again: "Is everything predetermined?" *Assuming that a TOE exists*, Hawking (1993, p. 139) finally says: "The answer is yes, it is. But it might as well not be, because we can never know what is determined." Isn't this a beautiful example of a profound truth (of which the logical contrary is also a profound truth) in the sense of Bohr (sec. 2.5)?

A final, and presumably conclusive, argument against complete predetermination comes from the Heisenberg uncertainty relation. Since a TOE must comprise quantum theory, it must also contain this uncertainty relation which is an ineluctable consequence of quantum theory (sec. 3.5). Cf. also (Gell–Mann 1994) and (Weinberg 1993).

Relation to dialectics. As we have seen, a true TOE must *derive itself*. This is another instance of the important concept of *self–reference*: As we have seen at the end of sec. 2.5, the dialectic system of Hegel's logic provides such a self–reference or self–derivation; cf. also (Speiser 1952, p. 110). This is possible in Hegel's informal thinking, but not so readily in the formal structures of contemporary logic and mathematics. It would have to be a *self–referential deductive system* (Wheeler 1994, p. 309).

In the light of these remarks it seems unlikely that we shall ever have a true TOE in the strict sense of the contradictory requirements (1) to (5) given above. What can happen if no TOE exists at all? Hawking (1988, p. 166) gives two alternatives.

(1) There is no final theory of the universe, only an infinite sequence of theories which better and better describe the universe. This is the present view of those physicists who do not believe in a universal TOE.

(2) There is no theory of the universe. The physical laws are valid only to a certain amount of accuracy; below this accuracy, events occur in a spontaneous and random fashion. This would be a general instance of the dialectics of freedom and necessity (sec. 2.5); this possibility has also been envisaged by Whitehead (1933, Chapter VII). We are also reminded of Wheeler's "no laws" (sec. 6.5).

Bootstrapping. Remarkably enough, superstring theory seems in fact to be able to "derive itself" in the form of "bootstrapping", cf. end of sec. 2.5: provided the system of elementary particles (as predicted by superstring theory) *exists*, it *produces itself.* This is the *bootstrap principle* of self–consistency (Gell–Mann 1994, Chapters 10 and 14); cf. also (Capra 1976, Chapter 18).

Logical singularities

Self–reference and self–derivation may lead to logical singularities. Examples have been given at the end of sec. 5.3 (especially Examples 1 to 3).

Logical singularities occur if the observer (or theoretician) A
(α) coincides with the object B or
(β) is part of the object B.

Case (α) has been seen to apply Plotin's definition of human thinking: *The thinking thinks the thinking* (sec. 2.5). Case (β) corresponds to B being the whole universe, of which A necessarily is a part (from Kant's antinomies to TOE). It is illustrated by Fig. 6.4.

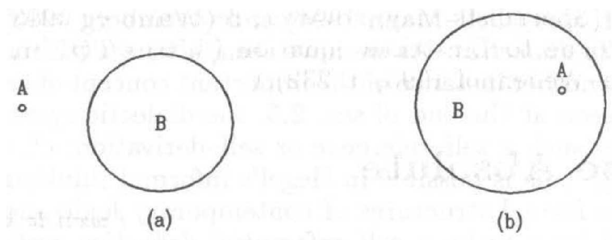


Figure 6.4: Subject A lies (a) outside object B, and (b) inside universe B

(*For specialists only.*)A simple model for such a logical singularity is provided by the mathematical singularity of the gravitational potential of an extended body B, well known to physicists and geodesists. The Newtonian potential V at point A is

$$V(A) = G \iiint_B \frac{\rho}{l} \, dv \quad , \tag{6.2}$$

where G is the gravitational constant, ρ is the density, dv is the volume element, and l denotes the distance of dv from A; the integral is taken over the whole body B (Fig. 6.5). Since in case (b), A may coincide with dv so that $l = 0$, formula (6.2) has a singularity at $l = 0$ for an internal point A, whereas it is regular for an external point, case (a).

In fact, both cases have quite a different behavior:

(a) $$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0 \qquad \text{Laplace's equation} \quad ,$$

(b) $$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -4\pi G\rho \qquad \text{Poisson's equation} \quad .$$

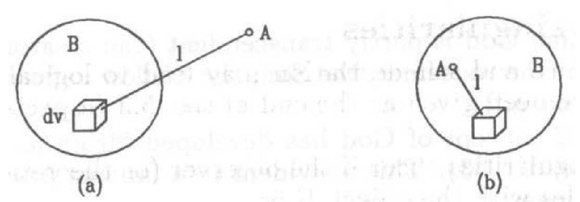


Figure 6.5: The Newtonian potential (a) for an external point A, and (b) for an internal point A.

This model comes entirely natural to geodesists and mathematical physicists, so it is given here. Readers from other disciplines can safely disregard it. (Specialists may, however, recognize equation (b) as the classical analogue to Einstein's equation (3.55) on p. 96, cf. (Moritz and Hofmann–Wellenhof 1993, p. 232).)

6.7 The Absolute

I want to know how God created this world.

Albert Einstein

Almost all metaphysical systems contain the concept of an Absolute, from Plato to Whitehead. He may be the creator of the world, according to the Jewish–Christian–Muslim tradition, he may be the Prime Mover and the supreme "Final Cause" of Aristotle and Teilhard de Chardin, he may be the metaphysical link between man and nature of Descartes, the Supreme Monad of Leibniz, the great Mechanical Engineer of Newton, the supreme Moralist of Kant, the Absolute Idea of Fichte and Hegel; or the origin of order, structure and harmony, the Supreme Mathematician of Plato, Kepler, Whitehead, and Hawking (1988, concluding sentence).

Generally we may distinguish in God

– *transcendence*, and

– *immanence*.

A *transcendent* God is fully above or outside the world (Jahwe, Allah). An *immanent* God is fully inside the world; this is the doctrine of *pantheism* (Spinoza, Hegel, dialectical materialism: matter is the Absolute, see above).

The Christian God is partly transcendent (the creator, the Father) and immanent (the redeemer, the Son, the "fellow sufferer who understands" (Whitehead)).

The biblical concept of God has developed "from force to persuasion" (Whitehead 1933). This is evident in a famous passage from the 1st Book of Kings:

> The LORD said, "Go out and stand on the mountain in the presence of the LORD, for the LORD is about to pass by." Then a great and powerful wind tore the mountains apart and shattered the rocks before the LORD, but the LORD was not in the wind. After the wind there was an earthquake, but the LORD was not in the earthquake. After the earthquake came a fire, but the LORD was not in the fire. And after the fire came a gentle whispering breeze. When Elijah heard it, he pulled his cloak over his face and went out and stood at the mouth of the cave.

Whitehead (1929, Part V, Chapter II) gives the following beautiful antitheses:

> It is as true to say that God is one and the World many, as that the World is one and God many.
>
> It is as true to say that the World is immanent in God, as that God is immanent in the World.
>
> It is as true to say that God transcends the World, as that the World transcends God.

These are perhaps the finest examples of "profound truths" in the sense of Niels Bohr. As a matter of fact, it is an established tradition in mysticism to speak of God in antitheses (cf. Huxley 1970). This is comparable to the paradoxes of the Infinite as outlined in sec. 2.3.

Curiously enough, *bootstrapping* (secs. 2.5 and 6.6) occurs also here, namely in the so–called *ontological argument* for the existence of God. It goes back to medieval philosophy, to Anselm of Canterbury (sec. 5.4). With considerable oversimplification, it may be formulated as follows: By definition, God is absolutely perfect in every respect. Since an existing being is more perfect than a non–existing one, absolute perfection entails existence. In this way, profanely speaking, God "bootstraps" Himself into existence.

The ontological argument has a colorful history. It was rejected already by St. Thomas Aquinas but revived by Descartes, Spinoza and Leibniz, rejected by Kant, revived by Fichte and Hegel, and rejected again by Schopenhauer.

A modern analysis was performed by Findley (1963, Introduction) and Ch. Hartshorne: If God's existence is any way *possible*, then it is also *certain* that God exists. But lest believers rejoice too early, the argument can be inverted: if God's nonexistence is in any way *possible*, it is *certain* that God does not exist!

This is indeed a logically fascinating paradox of the Infinite, and the little book (Plantinga 1965) is as intellectually rewarding as (Nagel and Newman 1958).

So far we have considered the concept, or concepts, of God mainly from a philosophical and logical point of view. So far it has been independent of revealed religion. This distinction between philosophy and theology (religion) was made already by St. Thomas Aquinas. An excellent, comprehensive and objective treatment of both points of view (philosophy and religion) was given by Küng (1978); an outstanding concise, clear and understandable summary is found in the last chapter of (Bocheński 1959).

Science and religion. It is a commonplace that science and religion do not contradict each other because they are dealing with different subject matters. Nevertheless, they have a common boundary, and wars are frequently waged about boundaries.

What should be rejected is the doctrine of two separate truths, one for science and one for religion. In fact, the synthesis between science and religion is not a cheap compromise, but requires considerations on a rather high level.

Such high–level efforts can be found, for example, in the writings of Whitehead and of Weizsäcker, and on the Catholic side of Teilhard de Chardin (1955), and in particular of Karl Rahner. Cheap popularizations frequently do more harm than good. Tipler's (1994) ingenious and ingenuous attempt to develop a theology on the basis of physics is probably a singularity, although an interesting one. (Davies 1983) and (Ferguson 1994) are more realistic. (Kolakowski 1982) is a fine counterpoint to (Küng 1978) by a former Marxist.

A chief stumbling–block has been the theory of biological evolution. The theory that man has evolved from animal predecessors was considered incompatible with man's immortal soul and hence with Christianity. This apparent contradiction was used by both sides to fight the other, with most unfortunate results.

In fact, even if man has evolved in this way (which no serious scientist would doubt nowadays), man's intellect is such that he is clearly separated from the other animals. Thus, he seems to be the only animal capable of thinking about himself. This "thinking about thinking" is *self–consciousness*, cf. (Eccles 1989).

Another really stupid stumbling–block has been the apparent contradiction between the creation in seven days according to the Bible and the history of evolution which took billions of years. Most unfortunately, theologians did not do their homework well: had they read St. Augustine, they would have found that the biblical days should not be taken literally: "With the Lord, a day is like a thousand years, and a thousand years is like a day" (2nd Letter of St. Peter). In fact, thus interpreted, we have a remarkable similarity of present–day scientific thinking and the biblical report:

BIBLE	SCIENCE
creation	"big bang"
light (1st day)	energy
Heaven and Earth (2nd day)	stars, planets, Earth
oceans,	oceans,
vegetation (3rd day)	vegetation
Sun and Moon, day	Earth rotation,
and night (4th day)	formation of the Moon
fish and birds (5th day)	animals
mammals, man (6th day)	mammals, man

With the possible exception of the 4th day, the order is almost exact; cf. (Weizsäcker 1973).

As we believe today, evolution is not restricted to biology, but started right after the "big bang". First hydrogen was formed, then helium was synthesized, and heavier elements followed. Galaxies and stars were formed. Stellar evolution is now standard in astronomy. The evolution of the planetary system was discussed already by Laplace and Kant.

The history of mankind and the story of the bible fit beautifully into this picture of a thoroughly evolutionary universe. Christianity was always proud of its historical evolution, from the first man ("Adam" simply means "man"!) to Moses, King David, and Jesus. Not to see that this historical development fits perfectly into general evolution seems, in hindsight, an almost unbelievable shortsightedness. Teilhard de Chardin (1955) was one of the first Catholics to work this into a consistent picture, and he had to pay for it. Nowadays, of course, this is no problem any more.

Can science help prove the *existence of God*? When Napoleon asked Laplace why he did not mention God in his famous book on celestial mechanics, Laplace answered: "Sire, I do not need this hypothesis". Of course, he was right from the point of view of classical mechanics: God may have installed the giant mechanical clockwork of the universe, but since then it continued to run automatically according to the rigorous laws of mechanics. Divine interference was no longer needed *in classical mechanics*.

Other scientists and philosophers thought differently: Kepler admired the divine harmony of the universe, and so did Newton. Leibniz tried hard to reconcile science and religion. Kant deduced the existence of God and the immortality of the soul from the moral law to which all men are subject. Goethe, following Spinoza, sought God in nature and art, Einstein in the wonderful mathematical laws of nature ("Science without religion is lame, religion without science is blind" (Einstein 1954, p. 46). Heisenberg and Weizsäcker are Evangelic Christians.

Also to me it is impossible to believe that a beautiful mountain flower or Riemann's formula for the distribution of prime numbers are pure products of "chance" or "natural selection".

Biologists, discussing evolution, frequently speak of the "creativity" or the "inventiveness" of evolution. The physicist Freeman Dyson (quoted from (Kaku 1994, p. 258)) writes:

> As we look out into the Universe and identify the many accidents of physics and astronomy that have worked together to our benefit, it almost seems as if the Universe must in some sense have known that we were coming.

Here obviously to evolution and to the Universe we find ascribed properties of the Absolute …

It is generally believed that the increasing complexity of evolution implies an increase of information (or a decrease of entropy). Does this imply a continuous supply of information "from the outside", some kind of "continuous creation"? Tresmontant (1976) and others think so.

The old clash between religion and science was largely due to Whitehead's "fallacy of misplaced concreteness". The exaggerated "spiritualism" of religion seemed to be incompatible with the exaggerated "materialism" of science. As we have seen, materialism no longer reigns supreme in modern physics. In relativity and quantum theory, science has proved to be infinitely more complex and counterintuitive than the most daring speculations of theologians and metaphysicists.

In fact, the main importance of modern science for religious thinking may be the fact that it has provided highly counterintuitive and strange models which make the old difficulties of religion (miracles etc.) look harmless in comparison. No one any longer believes that God really stopped the movement of the Sun in order to permit Joshua to win his battle against the Amorites in daylight: the bible is a religious work and not a textbook on Earth rotation! Also the allegoric interpretation of Scripture, already advocated by St. Augustine(!), is now being generally accepted in theology.

There is one real contribution of the hard theological work of the Fathers of the Church to contemporary thinking: the theory of the Trinity. Plotinus' statement "The thinking thinks the thinking" has been applied to God as follows: The thinking (1st person) thinks (3rd person) the thinking (2nd person). The Holy Trinity served not only as a basic model for mutual immanence which plays a fundamental role in the theory of internal relations (Whitehead 1933, Chapter X, sec. IV; 1925, Chapter XI). It also provided the prototype of a dialectic triad: thesis, antithesis, and synthesis.

Questions of the purpose of man, of the meaning of existence, and of ethics and care for coming generations have been thought to be beyond science. Now, in view of the nuclear threat, the destruction of the environment, and the possibility of genetic manipulation we start to think differently. I do not feel qualified to treat these supremely important questions here and refer the reader to the enormous literature, e.g. (Fischer 1993) and (Koltermann 1994).

6.8 Pluralism

Die Wahrheit ist symphonisch.
Hans Urs von Balthasar

Scientific pluralism is the peaceful coexistence of (mildly) conflicting theories such as classical mechanics, relativity and quantum theory, as well as of their interpretations, which are particularly controversial in quantum mechanics (sec. 3.5). Nevertheless we may speak of a general consensus between scientists.

The first impression on reading philosophy, however, is a bewildering number of apparently contradictory philosophical systems and "isms". There is practically nothing in common, say, between Heidegger and Russell. Which of these systems is true?

The answer is: none and all (another "profound truth" in the sense of Bohr!). In fact, every reasonable system of philosophy offers an aspect or perspective of reality, a "theory" in the sense of Bohm: a way of looking at our world. In Greek "*theoria*" means a way of looking, a perspective, as we have seen in sec. 6.5. Reality is far too complicated to fit into one single system, even too complicated to fit into a common physical theory.

A high mountain looks different as we regard it from north, south, east, or west. Which one is the "correct" perspective? Is the view from north "true" and the view from east "false"? Should the observer from south critize the viewer from the west?

In philosophy, however, it is customary for the adherents of one system to critize their colleagues from the other systems. Such criticism is necessary in order to make the reasoning logically stringent and convincing. It should be an objective and fair criticism, preceded by intense and sympathetic study of the other's *complete* reasoning.

Superficial criticism based on sentence–by–sentence refutation is easy and cheap. It is not difficult to condemn existentialism on the basis of logical positivism and vice versa. A particularly shocking misunderstanding regarding Fichte was mentioned in sec. 5.4.

It is quite improbable that great philosophical systems are completely erroneous. They at least contain some important truth or insight.

My general experience with philosophical criticism is that positive opinions of philosophers about other philosophers seem usually right, and that negative judgments of philosophers about other philosophers seem usually wrong. Why? I don't know, it might be related to human psychology.

All this speaks in favor of tolerance and scientific and philosophical pluralism. The plurality of philosophical systems is as necessary for human understanding as the plurality of musical instruments is for an orchestra. Figuratively speaking, Plato's violin is, of course, omnipresent. But Fichte's flute and Hegel's trumpet add color, and the cello of Kant and the contrabass of logical positivism provide a reliable basis. The exotic instruments of Gödel and Wittgenstein are used if the orchestra is to play Schönberg and Webern ...

Thus Hans Urs von Balthasar says (see the motto of this section):

"TRUTH IS SYMPHONIC".

Selected additional reading

It is quite clear to me that you will probably neither be able to read all the books given, nor even wish to do so. It is a rich menu, from which you are invited to select at your convenience.

You are asking for a suggestion? If you can read only *one* book, I suggest (Davies 1988), if you have time for two books, take also (Whitehead 1925), and as a third, (Penrose 1989), which is tougher but extremely rewarding. My next candidates on the reading list would be (Hofstadter 1979), (Cohen and Stewart 1994), and (Popper 1982). Further recommendations are found in the text of the book.

A triad of very readable and mutually complementary introductory booklets on "pure" philosophy is (Bocheński 1959), (Jaspers 1953) and (Russell 1912).

As far as available to the author, book titles have been given in English. Several German and French books are available in an English translation (and vice versa).

Abraham R H and Shaw C D (1992) Dynamics: The Geometry of Behavior, 2nd ed, Addison–Wesley, Redwood City, Cal.

Anger G, Gorenflo R, Jochmann H, Moritz H, and Webers W eds (1993) Inverse Problems: Principles and Applications in Geophysics, Technology, and Medicine, Akademie Verlag, Berlin (mathematical).

Barrow J D (1988) The World Within the World, Oxford Univ. Press.

Barrow J D (1991) Theories of Everything, Oxford Univ. Press.

Barrow J D (1992) Pi in the Sky, Oxford Univ. Press.

Blokhintsev D I (1968) The Philosophy of Quantum Mechanics, Reidel, Dordrecht, Holland (mathematical but also of general interest).

Bocheński J M (1959) Wege zum philosophischen Denken, Herder, Freiburg.

Bohm D (1980) Wholeness and the Implicate Order, Routledge, London.

Bohr N (1934) Atomic Theory and the Description of Nature, Cambridge Univ. Press.

Bohr N (1958) Atomic Physics and Human Knowledge, Wiley, New York.

Bohr N (1963) Essays 1958–1962 on Atomic Physics and Human Knowledge, Wiley, London. (All three volumes reprinted 1987 by Ox Bow Press, Woodbridge, Connecticut.)

Briggs J (1992) Fractals: the Patterns of Chaos, Simon and Schuster, New York.
Čapek M ed (1976) The Concepts of Space and Time, Reidel, Dordrecht, Holland.
Capra F (1975) The Tao of Physics, Wildwood House, London.
Carnap R (1950) Logical Foundations of Probability, Chicago Univ. Press (for specialists).
Carnap R (1958) Introduction to Symbolic Logic and Its Applications, Dover, New York (for specialists).
Carnap R (1966) Philosophical Foundations of Physics, Basic Books, New York.
Churchland P M (1988) Matter and Consciousness, 2nd ed., MIT Press.
Cohen L J (1986) The Dialogue of Reason, Oxford Univ. Press.
Cohen L J (1989) An Introduction to the Philosophy of Induction and Probability, Oxford Univ. Press.
Cohen J and Stewart I (1994) The Collapse of Chaos: Discovering Simplicity in a Complex World, Viking–Penguin, New York.
Copleston F (1946) A History of Philosophy, 9 vols., Burns and Oates, London.
Davies P (1983) God and the New Physics, Simon and Schuster, New York.
Davies P (1984) Superforce: The Search for a Great Unified Theory of Nature, Simon and Schuster, New York.
Davies P (1988) The Cosmic Blueprint, Simon and Schuster, New York.
Davies P (1992) The Mind of God: The Scientific Basis for a Rational World, Simon and Schuster, New York.
Davies P and Brown J R (1986) The Ghost in the Atom, Cambridge Univ. Press.
Davies P and Brown J R (1988) Superstrings: A Theory of Everything? Cambridge Univ. Press.
Ditfurth H von (1976) Der Geist fiel nicht vom Himmel, Hoffmann und Campe, Hamburg.
Duncan R and Weston–Smith M eds (1977) The Encyclopedia of Ignorance, Pergamon, Oxford.
Eccles J C (1973) The Understanding of the Brain, McGraw–Hill, New York.
Eccles J C (1989) Evolution of the Brain: Creation of the Self, Routledge, London.
Eccles J C (1994) How the Self Controls Its Brain, Springer, Berlin.
Eddington A (1939) The Philosophy of Physical Science, Cambridge Univ. Press.
Eddington A (1959) New Pathways in Science, Ann Arbor Paperback, Univ. of Michigan Press.
Edelman G M (1989) The Remembered Present: A Biological Theory of Consciousness, Basic Books, New York.

Edey M A and Johanson D C (1989) Blueprints: Solving the Mystery of Evolution, Penguin Books, New York.

Eigen M (1987) Stufen zum Leben, Piper, München.

Eigen M and Winkler R (1975) Das Spiel, Piper, München.

Einstein A (1954) Ideas and Opinions, Crown, New York.

Einstein A (1957) Mein Weltbild, Ullstein, Berlin.

Ferguson K (1994) The Fire in the Equations, Bantam Press, London.

Findley J N (1958) Hegel: A Re–examination, Allen and Unwin, London.

Findley J N (1963) Language, Mind and Value, Allen and Unwin, London.

Fischer H (1993) Plädoyer für eine sanfte Chemie, C.F. Müller, Karlsruhe.

Folse H J (1985) The Philosophy of Niels Bohr, North Holland, Amsterdam.

Gardner M (1982) The Ambidextrous Universe, 2nd ed, Penguin, London.

Gell–Mann M (1994) The Quark and the Jaguar: Adventures in the Simple and the Complex, Freeman, New York.

Gierer A (1985) Die Physik, das Leben und die Seele, Piper, München.

Gilson E (1972) L'être et l'essence, 2nd ed., J. Vrin, Paris.

Glass L and Mackey M C (1988) From Clocks to Chaos, Princeton Univ. Press.

Gleick J (1988) Chaos: Making a New Science, Cardinal–Penguin, London.

Globus G G, Maxwell G, and Savodnik I eds (1976) Consciousness and the Brain, Plenum Press, New York.

Gnedenko B V and Khinchin A Ya (1962) An Elementary Introduction to the Theory of Probability, Dover, New York.

Götschl J ed (1992) Erwin Schrödinger's World View, Kluwer, Dordrecht, Holland.

Gribbin J (1984) In Search of Schrödinger's Cat, Wildwood House, London.

Gulyga A (1990) Die klassische deutsche Philosophie, Reclam, Leipzig.

Hadamard J (1945) The Psychology of Invention in the Mathematical Field, Princeton Univ. Press.

Haken H (1981) Erfolgsgeheimnisse der Natur: Synergetik, 2nd ed., Deutsche Verlags–Anstalt, Stuttgart.

Haldane J B S (1939) The Marxist Philosophy and the Sciences, reprint by Books for Libraries Press, Freeport, New York.

Halmos P R (1960) Naive Set Theory, Van Nostrand, Princeton (mathematical).

Hartmann N (1960) Die Philosophie des deutschen Idealismus, 2nd ed., W. de Gruyter, Berlin.

Hartmann N (1965) Grundzüge einer Metaphysik der Erkenntnis, 5th ed., W. de Gruyter, Berlin.

Havemann R (1964) Dialektik ohne Dogma? Naturwissenschaft und Weltanschauung, Rowohlt, Hamburg.

Hawking S (1988) A Brief History of Time, Bantam Press, London.

Hawking S (1993) Black Holes and Baby Universes and Other Essays, Bantam Press, London.

Hawking S W and Ellis G F R (1973) The Large Scale Structure of Space Time, Cambridge Univ. Press (mathematical).

Heintel E (1990) Die Stellung der Philosophie in der "Universitas Litterarum", Publishing House of the Austrian Academy of Sciences, Wien.

Heisenberg W (1955) Das Naturbild der heutigen Physik, Rowohlt, Hamburg.

Heisenberg W (1958) Physics and Philosophy, Harper & Row, New York.

Heisenberg W (1973) Der Teil und das Ganze, Piper, München.

Hempel C G (1966) Philosophy of Natural Science, Prentice–Hall, Englewood Cliffs, N.J.

Hofstadter D R (1979) Gödel, Escher, Bach: an Eternal Golden Braid, Basic Books, New York.

Hofstadter D R (1985) Metamagical Themas: Questing for the Essence of Mind and Pattern, Basic Books, New York.

Hofstadter D R and Dennett D C, eds (1981) The Mind's I, Basic Books, New York.

Holzmüller W (1984) Information in Biological Systems: The Role of Macromolecules, Cambridge Univ. Press.

Hubel D H (1988) Eye, Brain, and Vision, Scientific American Library, New York.

Huxley A (1970) The Perennial Philosophy, Harper & Row, New York.

Jantsch E (1980) The Self–Organizing Universe, Pergamon, Oxford.

Jaspers K (1953) Einführung in die Philosophie, Piper, München.

Jaspers K (1962–1993) The Great Philosophers, Harcourt Brace Jovanovich, New York.

Jeffreys H (1961) Theory of Probability, 3rd ed., Oxford Univ. Press (mathematical).

Jeffreys H (1973) Scientific Inference, 3rd ed., Cambridge Univ. Press.

Jordan P (1968) Der Naturwissenschaftler vor der religiösen Frage, 5th ed., Stalling, Oldenburg.

Kaku M (1994) Hyperspace: A Scientific Odyssey Through Parallel Universes, Time Warps, and the 10th Dimension, Oxford Univ. Press.

Kauffman S A (1993) The Origins of Order: Self–Organization and Selection in Evolution, Oxford Univ. Press.

Kohonen T (1988) Self–Organization and Associative Memory, 2nd ed., Springer, Berlin.

Kolakowski L (1982) Religion — If there is no God, Fontana Paperbacks, London.

Koltermann R (1994) Grundzüge der modernen Naturphilosophie, Knecht, Frankfurt a. M.

Kosko B (1993) Fuzzy Thinking, Hyperion, New York.

Kuhn T S (1970) The Structure of Scientific Revolutions, 2nd ed., Chicago Univ. Press.
Küng H (1978) Existiert Gott? Piper, München.
Küppers B O ed (1987) Leben = Physik + Chemie? Piper, München.
Kuznecov B G (1981) Philosophie – Mathematik – Physik, Akademie–Verlag, Berlin.
Lanczos C (1965) Albert Einstein and the Cosmic World Order, Wiley–Interscience, New York.
Lanczos C (1970) Space Through the Ages, Academic Press, London (rather mathematical).
Leff H S and Rex A F eds (1990) Maxwell's Demon, Adam Hilger, Bristol.
Lewin R (1992) Complexity: Life at the Edge of Chaos, Macmillan, New York.
Lindsay R B and Margenau H (1957) Foundations of Physics, Dover Publications, New York (rather mathematical).
Lockwood M (1989) Mind, Brain and the Quantum, Basil Blackwell, Oxford.
Lorenz E N (1993) The Essence of Chaos, Univ. of Washington Press, Seattle.
Lorenz K (1973) Die Rückseite des Spiegels, Piper, München; English translation: Behind the Mirror, Methuen, London, 1977.
Lucas J R (1970) The Freedom of the Will, Oxford Univ. Press.
Lucas J R (1973) A Treatise on Time and Space, Methuen, London.
Margenau H (1950) The Nature of Physical Reality, McGraw–Hill, New York.
Margenau H (1984) The Miracle of Existence, Ox Bow Press, Woodbridge, Connecticut.
Maturana H R and Varela F J (1987) Der Baum der Erkenntnis, Scherz, Bern.
Mayer–Kuckuk T (1989) Der gebrochene Spiegel, Birkhäuser, Basel (on symmetry, symmetry breaking, and order in nature).
Miller D ed (1985) Popper Selections, Princeton Univ. Press.
Misner C W, Thorne K S, and Wheeler J A (1973) Gravitation, Freeman, San Francisco (mathematical).
Monod J (1970) Le hasard et la necessité, Éditions du Seuil, Paris.
Moritz H and Hofmann–Wellenhof B (1993) Geometry, Relativity, Geodesy, Wichmann, Karlsruhe (mathematical).
Moser F (1989) Bewusstsein in Raum und Zeit, Leykam, Graz.
Müller G E (1974) Dialektische Philosophie, A. Francke Verlag, München.
Nagel E and Newman J R (1958) Gödel's Proof, New York Univ. Press.
Nahin P J (1993) Time Machines: Time Travel in Physics, Metaphysics, and Science Fiction, American Institute of Physics, New York.
Nicolis G and Prigogine I (1989) Exploring Complexity, W.H. Freeman, New York.

Oeser E and Seitelberger F (1988) Gehirn, Bewusstsein und Erkenntnis, Wissenschaftliche Buchgesellschaft, Darmstadt.

Ornstein R and Thompson R F (1984) The Amazing Brain, Houghton Mifflin, Boston.

Penrose R (1989) The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics, Oxford Univ. Press.

Perry J and Bratman M eds (1993) Introduction to Philosophy: Classical and Contemporary Readings, 2nd ed., Oxford Univ. Press.

Piaget J (1970) L'épistémologie génétique, Presses Universitaires de France, Paris.

Plantinga A ed (1965) The Ontological Argument, Doubleday Anchor Books, Garden City, New York.

Poincaré H (1902) La science et l'hypothèse, Flammarion, Paris; reprint: Science and Hypothesis, Dover, New York.

Poincaré H (1908) Science et méthode, Flammarion, Paris; reprint: Science and Method, Dover, New York.

Poincaré H (1958) The Value of Science, Dover, New York (reprint).

Poincaré H (1963) Mathematics and Science, Dover, New York (reprint).

Popper K R (1977) The Logic of Scientific Discovery, 9th ed., Hutchinson, London.

Popper K R (1979) Objective Knowledge: An Evolutionary Approach, 2nd ed., Oxford Univ. Press.

Popper K R (1982) The Open Universe: An Argument for Indeterminism, Routledge, London.

Popper K R and Eccles J C (1977) The Self and Its Brain, Springer, Berlin.

Prigogine I and Stengers I (1984) Order out of Chaos, Bantam Books, New York.

Quine W V O (1961) From a Logical Point of View, 2nd ed., Harvard Univ. Press.

Reichenbach H (1957) The Philosophy of Space and Time, Dover, New York.

Rosenthal D M ed (1991) The Nature of Mind, Oxford Univ. Press.

Rucker R (1982) Infinity and the Mind, Birkhäuser, Boston.

Rucker R (1984) The Fourth Dimension, Houghton Mifflin, Boston.

Ruhla C (1992) The Physics of Chance, Oxford Univ. Press.

Russell B (1912) The Problems of Philosophy, William and Norgate, London.

Russell B (1927) An Outline of Philosophy, Allen and Unwin, London.

Russell B (1929) Our Knowledge of the External World, Allen and Unwin, London.

Russell B (1945) A History of Western Philosophy, Allen and Unwin, London.

Russell B (1948) Human Knowledge: Its Scope and Limits, Allen and Unwin, London.

Schilpp P A (1949) Albert Einstein: Philosopher – Scientist, 2 vols., Harper and Brothers, New York.

Schrödinger E (1944) What is Life? Cambridge Univ. Press.

Schrödinger E (1954) Nature and the Greeks, Cambridge Univ. Press.

Schrödinger E (1958) Mind and Matter, Cambridge Univ. Press.

Sheldrake A R (1981) A New Science of Life, Blond & Briggs, London.

Singer W ed (1990) Gehirn und Kognition, Spektrum Akademischer Verlag, Heidelberg.

Singer W ed (1994) Gehirn und Bewusstsein, Spektrum Akademischer Verlag, Heidelberg.

Smart J J C ed (1964) Problems of Space and Time, Macmillan, New York.

Smullyan R M (1977) The Tao is Silent, Harper and Row, San Francisco.

Smullyan R M (1980) This Book Needs No Title: A Budget of Living Paradoxes. Prentice–Hall, New York.

Speiser A (1952) Elemente der Philosophie und der Mathematik, Birkhäuser, Basel.

Speiser A (1955) Die geistige Arbeit, Birkhäuser, Basel.

Speiser A (1959) Ein Parmenides–Kommentar, 2nd ed, Koehler, Stuttgart.

Spies M (1993) Unsicheres Wissen, Spektrum Akademischer Verlag, Heidelberg.

Squires E (1990) Conscious Mind in the Physical World, Adam Hilger, Bristol.

Stapp H P (1993) Mind, Matter, and Quantum Mechanics, Springer, Berlin.

Stegmüller W (1969) Metaphysik, Skepsis, Wissenschaft, 2nd ed., Springer, Berlin.

Stewart I (1990) "Does God Play Dice?", Penguin Books, London.

Teilhard de Chardin P (1955) Le phénomène humain, Éditions du Seuil, Paris.

Thom R (1975) Structural Stability and Morphogenesis, Benjamin, Reading, Massachusetts.

Thompson R F (1985) The Brain, Freeman, New York.

Tipler F J (1994) The Physics of Immortality, Doubleday, New York.

Treder H J (1974) Philosophische Probleme des physikalischen Raumes, Akademie–Verlag, Berlin.

Tresmontant C (1971) Le problème de l'âme, Éditions du Seuil, Paris.

Tresmontant C (1976) Sciences de l'univers et problèmes métaphysiques, Éditions du Seuil, Paris.

Vollmer G (1990) Evolutionäre Erkenntnistheorie, 5th ed, Hirzel, Stuttgart.

Waldrop M M (1992) Complexity: the Emerging Science at the Edge of Order and Chaos, Simon & Schuster, New York.

Weinberg S (1977) The First Three Minutes, Andre Deutsch, London.

Weinberg S (1993) The Dream of a Final Theory, Vintage, London.

Weizsäcker C F von (1970) Zum Weltbild der Physik, 11th ed., Hirzel, Stuttgart.

Weizsäcker C F von (1971) Die Einheit der Natur, Hanser, München.

Weizsäcker C F von (1973) Die Tragweite der Wissenschaft, 4th ed., Hirzel, Stuttgart.

Weizsäcker C F von (1977) Der Garten des Menschlichen, Hanser, München.

Weizsäcker C F von (1985) Aufbau der Physik, Hanser, München.

Weizsäcker C F von (1992) Zeit und Wissen, Hanser, München.

Weyl H (1949) Philosophy of Mathematics and Natural Science, Princeton Univ. Press.

Weyl H (1952) Symmetry, Princeton Univ. Press.

Wheeler J A (1994) At Home in the Universe. American Institute of Physics, Woodbury, New York.

Whitehead A N (1925) Science and the Modern World, Macmillan, New York.

Whitehead A N (1929) Process and Reality, Macmillan, New York (difficult!).

Whitehead A N (1933) Adventures of Ideas, Macmillan, London.

Whitrow G J (1980) The Natural Philosophy of Time, 2nd ed., Oxford Univ. Press.

Wiener N (1961) Cybernetics, 2nd ed., Wiley, New York.

Will C M (1986) Was Einstein Right? Basic Books, New York.

Young J Z (1987) Philosophy and the Brain, Oxford Univ. Press.

Zurek W H (1990) Complexity, Entropy and the Physics of Information, Addison–Wesley, Redwood City, California.

Index