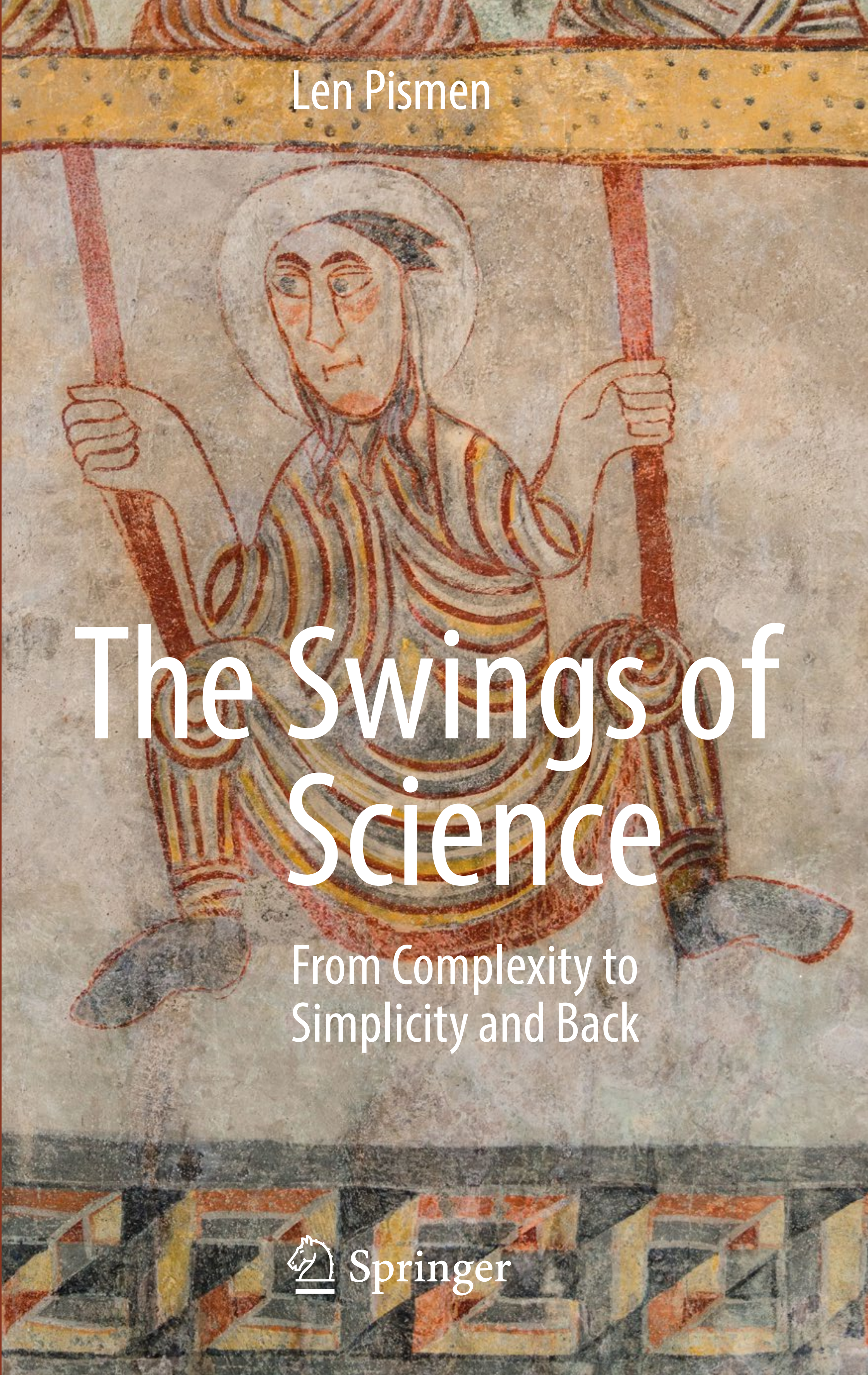
Len Pismen

The Swings of Science

From Complexity to Simplicity and Back



Springer

The Swings of Science

Len Pismen

The Swings of Science

From Complexity to Simplicity and Back

Springer

Len Pismen
Department of Chemical Engineering
Technion Israel Institute of Technology
Haifa, Israel

ISBN 978-3-319-99776-6 ISBN 978-3-319-99777-3 (eBook)
https://doi.org/10.1007/978-3-319-99777-3

Library of Congress Control Number: 2018955719

Caption: Romanesque fresco of St. Proculus in the church St. Proculus, 7th century, Naturns, South Tyrol, Italy. Notable is the wrong grip to the rope.
Credit: Dietrich Krieger.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book comes from an old man who has wandered among different branches of macroscopic physics (see http://pattern.technion.ac.il/), sometimes looking with envy at the lofty heights above and abysses below where “fundamental” physics either soars or penetrates. Most books being published nowadays for the inquisitive lay public are dedicated to these extremes, which, as we know, come together, as many extremes do. Although I cannot avoid them in the course of my narrative, in particular, because they are the parts that swing most amply, this is a book on “normal” science, about which Alfred Tennyson once said: “science moves, but slowly slowly, creeping on from point to point”, science closer to our scale, science that used to be fundamental once, and is still fundamental in unraveling all the complications left behind by those aspiring to reach the extremes. I shall talk not only about successes, but also about failures and controversies, which are unavoidable because science is done by humans.

The two seeds of this book are my old book in Russian (Pismen, 1973) and a lecture for an audience which theoretically had to encompass civil servants and bankers, but in fact consisted largely of people like myself, at the seminar organized by the Institut de l’École normale supérieure in 2009, and later published as a book chapter (Rubio et al., 2013). No crystal is structured, however, in the same way as the seed that helped it nucleate. This is not an objective account, but a personal story of science in time, and glimpses of my own life sometimes appear among more important topics.

I dedicate this book to Hannah Timerman, my granddaughter, still a toddler at the time of writing, with the hope that she re-reads it, or at least part of it, as an old woman at the turn of the twenty-second century, and contemplates how things have changed by then.

Haifa, Israel
2018

Len Pismen

Contents

Chapter 1
Swings Through the Ages

1.1 Complex or Simple?

There are many definitions of complexity, all of them inadequate. "Any definition of complexity is context dependent, even subjective" (Gell-Mann, 1994). We feel intuitively what is complex and what is simple – but perhaps it is an illusion? A coder's measure of complexity is the length of a program that can simulate the system. Is that any good? A Julia set is generated by a very short program which, when iterated, gives an infinitely intricate pattern on an ever refined scale – looking very complex to the eye (Fig. 1.1). Is $E = mc^2$ complex? It looks to be a very simple formula that every layperson now knows – so why did we need an Einstein to come up with it? – but its consequences may be very complex, even lethal (think of Hiroshima).

On the other hand, even a most astute hacker would get a headache trying to code the most primitive hominid band. With no program available, how could we decide when human society was more complex – now or twenty thousand years ago? Certainly, everything was simpler once when Adam was employed as a hunter, Eve as a gatherer, and Snake as a shaman, while now we have a banker, a garbage collector, a policeman, a social worker, a tennis player, a garage mechanic, and, and, and ..., and all sorts of laws, and all kinds of infrastructure, and scientific, pseudoscientific, and popular journals, and social media, and, and, and .... On the other hand, we all, or all those who enjoy or suffer this variety, live in the same global village, pass the same security gates to board the same planes, even speak the same mongrel English with different accents, while back then one could possibly walk the length of our weekend drive to encounter perhaps the same lifestyle, but a different language and different spirits and different legends and different gods.

It is still harder to decide what we actually prefer – simple or complex? Do we enjoy a simple formula or a complex picture that it generates? From the beginning of civilization, humans tried to introduce order into the infinite complexity of the surrounding world. This is what great ancient philosophers and great modern scientists aspired to do, from pre-Socratics to Albert Einstein, Paul Dirac, and Werner Heisenberg, vainly searching for a universal theory in their old age, to their living

L. Pismen, *The Swings of Science*,
https://doi.org/10.1007/978-3-319-99777-3_1

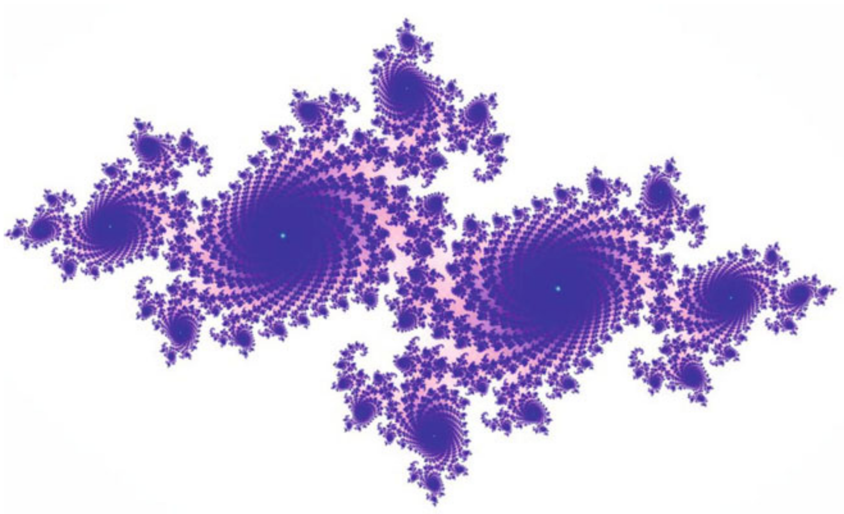


Fig. 1.1 Julia set

heirs still engaged in the same quest. In the words of one of them, "simplicity, properly understood, explains what it is that makes a good explanation deep, elegant, and beautiful" (Wilczek, 2012). On the other hand, totalitarian regimes of the 20th century tried to restrict the complexity of human thought to simple notions, be it class or racial struggle, and to impose uniform order. This was no innovation, as Plato aspired to the same in his *Republic*. The totalitarian order in its simplicity turns out, however, to be fragile and hard to sustain in the long run, as it gets overpowered by the ever present ambient complexity, as well as by the internal complexity of man, who never stops at simple explanations and simple rules. Simplicity is properly understood in conjunction with complexity. A universal theory breaks down in its applications and extensions to the infinite phenomenological variety in the midst of which we live and which we enjoy, and it is this complexity that makes it beautiful.

Boris Pasternak expressed this contradiction:

> We cannot avoid falling, as into heresy,
> into an unheard-of simplicity.
> But we will not be spared if we do not conceal it:
> people need it most of all,
> but they understand complexity better.

People need simplicity, to grasp the essence of things, maybe in a single moment of revelation – but they understand complexity better, having evolved to endure the challenges of a complex environment. This is the cause of eternal oscillations, from complexity to simplicity and back.

1.2 Swings of Religions

We can already see this swinging motion, from complexity to simplicity and back, in the evolution of religions, which historically served in many ways the same epistemological purpose that science is serving today (I shall say nothing about the other functions of religion). Animism was the "infant philosophy of mankind" (Tylor, 1871). There was a fairy in every tree, and a nymph in every stream, and one could talk to them. The world around was brought to life by its spirits, just as we were, and the common nature of spirits, inanimate, living, and dead, explained mysterious phenomena, like the travels of the soul while dreaming. As individual spirits coalesced into gods, like the spirits of trees into a god of the forest, their complexity became better organized, but these gods were still not so far removed from us.

> Some gods need not be praised,
> they are as equal with you,
> and with a careful hand
> you can rearrange them. (Osip Mandelstam)

There was an enormous variety of them, as each hunter–gatherer band had their own guardians and their own favorites.

The complex world was taken as is, without our analytical insights. The world was a world of miracles, both beneficial and adverse. We call a miracle something that is beautiful and unexpected, but we are not accustomed to miracles and we do not believe in them. Living in a world where miracles are common would be uncomfortable and dangerous, like living in a country with no laws. Magical shamanic rituals served both to federate a tribe and to obtain practical benefits by placating the spirits or gods in the mysterious irrational world. A rough analogy would be bribing officials to advance our aims or to avoid trouble. This helps to achieve practical benefits, while the unlawful world remains intact, miraculous, and mysterious. The

Fig. 1.2 The Magics

utility of rituals could even be tested experimentally, and gods could be punished if they failed to deliver the goods.

In modern terms, this is a "black box" approach, monitoring the reaction of a system to certain inputs without attempting to understand the internal mechanism. This is what you're doing when you press buttons on a hung laptop, trying to bring it back to life; this is the way the alchemist mixes up his concoctions in search of an elixir; this is how the statistician examines correlations; and it is even the way an engineer optimizes a process without the benefit of a mathematical model. Sometimes it even helps! The laptop starts working again, and the shaman, through incantations and herbs with their mixed psychological and physiological effects exorcises the illness with no less success than a certified doctor who is also unable to fully penetrate into the mechanism of the disease and its bearer. The consequences can also be dreadful, as in certain ill-fated experiments with the Chernobyl reactor. A market consultant operates not unlike a shaman, though with different incantations, and is remunerated as profusely. In any case, no lasting progress can be achieved in this way and the society relying on magic or black box methods is doomed to stagnation.

More order was introduced by ancient religions with their hierarchical pantheons and codified rituals. Gods, though powerful and immortal, were still humanlike, with their squabbles, rivalries, and love affairs. They still liked sacrifices, had their favorites, and interfered profusely in human affairs, as we know from reading Homer. Serving gods was always a prestigious and profitable occupation, and supreme rulers often doubled as supreme priests. Traditions consolidated into myths, and empires consolidated their pantheons, with gods serving as "geopolitical lubricants", redeeming the slaughter of conquest by the ensuing calm (Wright, 2009). The Romans considered the gods of subjugated peoples to be Roman gods by other names: it was a wise strategy retaining unity in diversity.

Variety and tolerance were incompatible with the monotheism gradually developing in Egypt and Mesopotamia and culminating in a unique God as the source and the cause of everything. The tetragrammaton, the unpronounceable sacred name, is the active form of the verb "to be", loosely translated as "he bringeth into existence". A direct line can be drawn from here to the dream of modern physics, "the theory of everything". When we now read the first chapter of the Book of Genesis, it sounds not unlike the theory of phase transitions in the early Universe, as explained to a

Fig. 1.3 *Left*: Zeus (Church, 1879). *Right*: Norse gods (Olaus Magnus, *Historia de gentibus septentrionalibus*)

pastoral tribe, followed by an account of the creation of living forms going in about the same order as in the theory of evolution.

On the other hand, the monotheism of Hebrew prophets, as well as other unified religio-philosophical systems born in the "axial age" in India, China, and Greece, brought forward ideals of a dedicated moral life and humanistic values (remaining, of course, ideals rather than reality). This was the period, centered around 500 BC, during which, according to Karl Jaspers (1953), "the spiritual foundations of humanity were laid simultaneously and independently". Spiritual ideals did not even require gods for their affirmation, as in Buddhism, Confucianism, or, for that matter, modern humanistic atheism. Adherence to a strict ritual made possible Jewish survival through two millennia, overcoming persecution and dispersion. Gothic cathedrals serving the Great God dwarfed all human constructions in the same way as CERN tunnels do nowadays – but rising to the skies rather than jostling through the bowels of the Earth.

Complexity found its way back when the unified systems were expanded and elaborated in diverse ways: in Judaism, by the sprawling commentaries of Talmudic scholars and eventually, back to magic, by Kabbalists and Hasidic mystics; in Christianity, by expanding the unique God to the Trinity and venerating a multitude of saints; in Buddhism, by the proliferation of schools and the cult of bodhisattvas. It was counteracted by a retrograde simplification in Islam, returning to the ethos of desert tribes, and in the various iconoclastic and puritanical movements throughout Christian history, and eventually in the totalitarian ideologies of the 20th century.

In our time, we see a proliferation of sectarian and esoteric beliefs, as well as all shades of traditional religions unrestricted, thanks to globalization, by geography or ethnicity. Science remains a thin crust over the molten lava of ignorance. Richard Feynman (1999), speaking at the Galileo Symposium in 1964, said that he would be ashamed to show Galileo our world today, which is so "actively, intensely unscientific". Why, Galileo would say, are there still astrologers when I proved that Jupiter is a ball with moons and not a god in the sky? Feynman continues to complain that telepathy and faith-healing are still alive, and in jest suggests trying to improve their performance by scientific methods. Society at large views science rather like a hog from a fable that does not care that the oak will wilt when its roots have been dug out, but only requires the oak's acorns to remain available. The fruits of science, from communication tools to weapons, are eagerly used by the same forces, from terrorists to preachers to pop stars, who undermine its roots.

1.3 Swings of Elements

In science, starting with its pre-scientific origins, the quest for simplicity is seen, first of all, in the quest for elementary entities in Nature. This was a fantastically deep simplifying concept: the entire diversity of objects and phenomena was envisaged as a combination of a few basic indivisible entities, as illustrated somewhat roughly by the artistic rendering in Fig. 1.4. The extreme unifying idea of Thales, not unlike monotheistic religious visions, was that water is the foundation of all; by

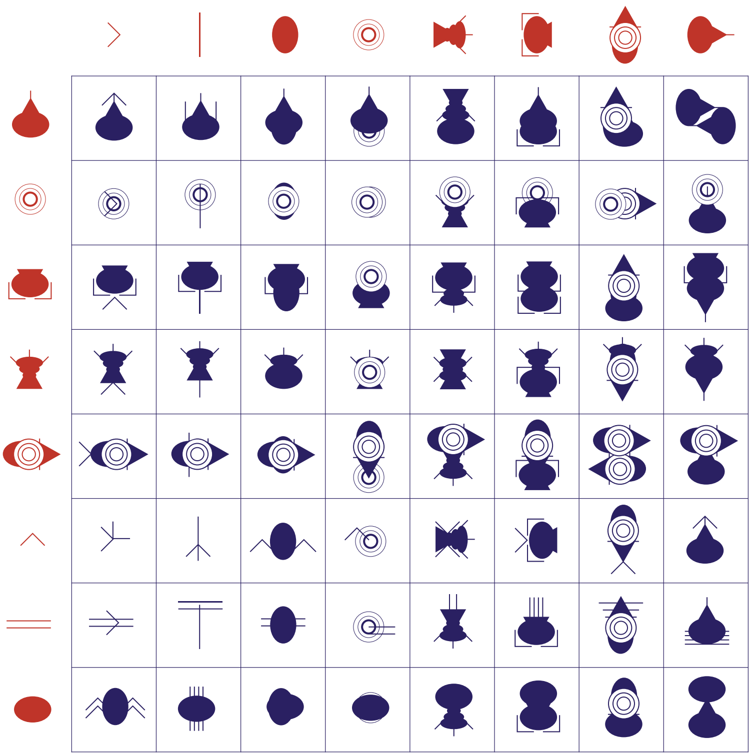


Fig. 1.4 Assembling elements into composite structures (Timerman and Timerman, 2002)

comparison, the symmetry-breaking transition in the first chapter of Genesis also involved waters, separated by a domain wall – the firmament of heaven.

This far-reaching universality was relaxed to four classical elements. Initially, the elements were images or attributes of the gods. Empedocles, probably influenced by the Babylonians and/or the Egyptians, associated the four elements – fire, air, water, earth – with the two wedded couples: Zeus and Hera, Persephone and Hades. The same elements, probably stemming from the same source, appear in the Buddha's teaching, while a somewhat different list of five elements is found in Indian Vedas and Chinese "states of being".

Divine associations were dropped by Aristotle, who also added the fifth element – aether, or quintessence, appearing in the Socratic dialog *Timaeus* as the most translucent kind of air breathed by the gods. The rational Aristotelian system dominated for almost two millennia. The elements were elegantly tied together by paired combinations of essential properties: dry fire and earth, wet air and water, hot fire and air, cold water and earth (Fig. 1.5). We could associate the four elements, not with the gods, but with the four states of matter: respectively, plasma, gas, liquid, and solid. Aether was resurrected in the 19th century as the space-filling medium car-

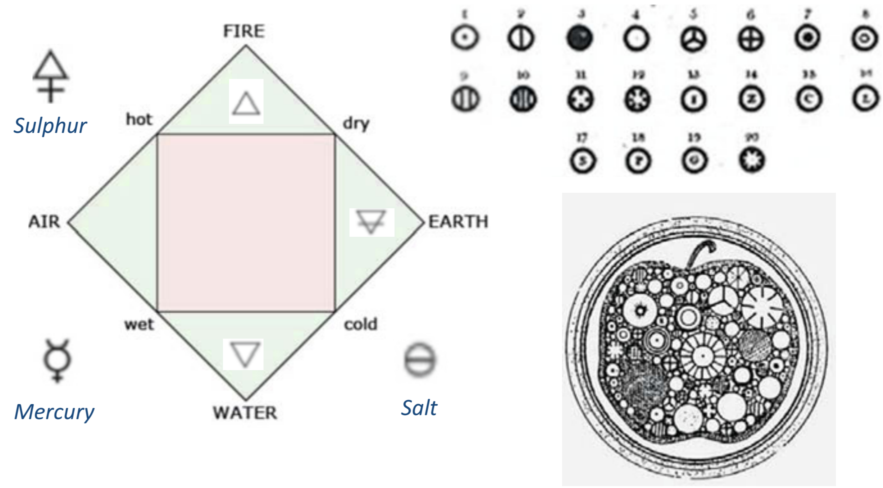


Fig. 1.5 *Left*: Aristotelian and medieval elements. *Upper right*: Dalton's elements. *Lower right*: symbolic representation of the composition of matter

rying electromagnetic waves (see Sect. 2.5), and quintessence as the hypothetical substance of dark energy at the turn of the 21st century.

The authority of Aristotle, endorsed by Thomas Aquinas, could not be challenged through the Middle Ages, but alchemists, more practically minded, were confused by its duality, which separated abstract qualities from their material carriers, but with neither appearing to be a primary principle. Jabir Ibn Hayyan added three material elements crucial for transmutation activities: sulphur, mercury, and salt (Burckhardt, 1967), the "*Tria Prima*" of European alchemists. These elements were not identical to the mundane substances bearing the same names. Thus, mercury was not just mercury, but philosopher's mercury, the embodiment of metallicity and mutability. The system of elements gradually became garbled and entangled, as Aristotle's authority would not allow it simply to be abolished, but only retrofitted by adding more elements.

In the new age, a completely different empirically based principle was suggested for identifying the elements: they were defined as substances that could not be split into simpler constituent parts. This was the basis of John Dalton's system of elements (Fig. 1.5); their number was not restricted, and, starting with the 33 elements listed by Antoine Lavoisier, this system kept expanding throughout the rational 19th century. "Chymistry" shed the Arab prefix, returning to the ancient Egyptian root, and the term "element" (Greek *stoicheion*) became the root of *stoichiometry*, introduced by Antoine Lavoisier to establish order in this rational chemistry (see Sect. 3.1).

The Greek philosophers would have been appalled by the disorganized multiplicity of elements. An early attempt to peer into a deeper level was undertaken by William Prout (1815), who observed that the atomic weights of the elements known at that time are close to integer multiples of the atomic weight of hydrogen, which appeared to suggest that the hydrogen atom was the only truly fundamental object. In retrospect, we can observe that the nuclei of all elements include hydrogen nuclei – protons – but, alas, there are also neutrons of about the same mass, something nobody would guess for another century. Naturally occurring elements are mixtures of isotopes containing different numbers of neutrons, and it was soon observed that the atomic weights of some elements did not comply with the integer rule. In the 19th century, experimental evidence trumped the beauty of the theory.

What was called "atomic weight" did not presume the actual existence of atoms; it was just a number inferred from the proportions of chemical species in different chemical reactions. The notion of elements was naturally compatible with the atomic hypothesis going back to Leucippus and Democritus in the fifth century BC: to each indivisible element corresponded an indivisible atom, and compound species, with their specific properties, could be assembled by combining atoms, as seen in the symbolic artist's rendering in Fig. 1.4. The concept of atoms could explain in a natural way why elements always react in ratios of whole numbers, but the existence of atoms had not yet been proven, and many prominent scientists, like Wilhelm Ostwald, Ernst Mach, and Dmitry Mendeleev, still believed in the infinite divisibility of matter at the turn of the 20th century, even after the discovery of radioactivity, which made atoms both real and not really atomic but divisible.

This was particularly ironic in the case of Mendeleev, who organized the multitude of elements known at the time into his famous Periodic Table. Mendeleev was fond of playing solitaire, and his greatest success in this game was setting up cards showing the various known elements in order of increasing atomic weight in such a way that elements with similar properties lay upon each other like playing cards in a set order. Such attempts had already been made, but Newlands (1864), when presenting his "octaves", had been ridiculed by someone who suggested that he set the elements in alphabetical order. Mendeleev's great solitaire game converged to his Periodic Table. But he had to cheat a bit. Would it be proper to put a knave on a king and skip the queen, or worse, interchange a queen and a knave to win the round? Mendeleev (1871) took the liberty of switching the places of some elements to fit the periodic changes in their chemical properties, and to leave empty cells for "queens", presumably still undiscovered in the chemical stack of cards (see Fig. 1.6).

Three missing elements fitting the table precisely were discovered in his lifetime: scandium (21), gallium (31), and germanium (32), while radioactive technetium (43), not occurring naturally, was isolated in 1937. Mendeleev's table also accommodated in a natural way the entire group of noble gases discovered by Lord Rayleigh and William Ramsay (1894).[1] Later corrections only inserted the rows of lanthanides and actinides, where single elements stood in the original table. Switching the positions appeared to be highly objectionable, since it was the atomic weight that was presumed to quantify the periodic changes. However, this move was vindicated soon after Mendeleev's death.

[1] Rayleigh and Ramsay received the 1904 Nobel Prize in Chemistry for the discovery of the inert gaseous elements in air and for determining their place in the periodic system, but Mendeleev, still alive, was overlooked, as was Lev Tolstoy.

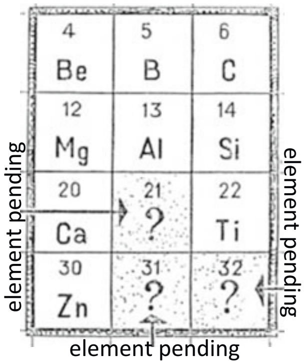


Fig. 1.6 Mendeleev's predictions

Soon after the neat arrangement of chemical elements had been achieved, atoms were proven to be both real and not elementary when Henri Becquerel (1896) and Pierre & Marie Curie discovered radioactivity (Curie et al, 1898) and Ernest Rutherford discovered atomic nuclei (Rutherford, 1911). Soon afterwards, Niels Bohr (1913) came up with his model of the atom (more on this in Sect. 6.1), which suggested that the spectral frequencies should be related to the charge of the nucleus. This model did not literally apply to heavy atoms with a large number of electrons, but the young Henry Moseley (1913) guessed that an approximate relation should hold, and found that the frequency of the main X-ray emission line correlated with the number of the element in Mendeleev's table, accounting for Mendeleev's transpositions and for as yet undiscovered elements. This was an amazing breakthrough: what was really important was not the atomic mass, carefully measured by 19th century chemists, but just the innocuous natural number assigned to the element in the Periodic Table. Moseley did not live to make more discoveries: a year later, he received a Turkish bullet at Gallipoli instead of a Nobel Prize[2].

It was now understood that the atom consisted of a nucleus with charge equal to its number in Mendeleev's table, surrounded by the matching number of electrons. Rutherford called the positively charged nuclear particles protons, and inferred the existence of neutral particles of comparable mass, viz., neutrons, to account for the difference between the atomic mass and the charge. By chance, a mixture of isotopes can be such that an element with a lower charge is heavier. Mendeleev could not have known anything like this, but dared and won.

A new cycle commenced in the history of the elements. Each time, elementary entities were given suitable names: when atoms (indivisible) turned out to be divisible, the next level entities were called elementary particles, only to lose that status in a mere half a century. The early 20th century was a moment of unheard-of simplicity. Only three elementary particles were needed: the electron, the proton, and the neutron. However, we were not long spared, for complexity kept creeping in: the relativistic quantum theory devised by Dirac (1928) required an antiparticle partner for each particle, and by the mid 20th century accelerators were producing a cornucopia of baryons and mesons, arbitrarily named and obeying no particular order. The very essence of elementarity appeared to be questionable, as parts sometimes turned out to be more massive than the whole, which was no wonder, as the equivalence of mass and energy had already been firmly established.

The next victory for simplicity came with the establishment of the Standard Model. The new set of elements were quarks and leptons, neatly organized in three generations (Fig. 1.7). Quarks, proposed by Murray Gell-Mann (1964), are an odd

---

[2] Freeman Dyson (1988) observed that "he died, so we would live". In World War II, promising scientists were not sent to fight but engaged intellectually, as Dyson was.

kind of animal that always sit in a cage of strongly interacting particles (hadrons, like the proton, neutron, and others), tied up by gluons, and never free to roam. Their family is bound by the symmetry of the "Eightfold Way" proposed by Gell-Mann (1962) and Yuval Ne'eman (1961). By this time, the nomenclature had become somewhat casual: Gell-Mann took the term "quark" from an obscure line in Joyce's *Finnegans Wake*. Moreover, quarks come in six flavors and three colors, terms that would better suit a bland German dairy product bearing the same name. The scheme of the Standard Model was subsequently enriched by incorporating the symmetry of weak interactions and the Higgs boson, responsible for the masses of the particles. All particles predicted theoretically to complete the scheme have since been discovered experimentally, with the Higgs, its long elusive cornerstone, coming last in 2012.

This unification is still considered to be unsatisfactory, as it contains an excessive number of parameters (masses and mixing angles), which, as adepts of the "Theory of Everything" will tell you, should come from a unification on a higher level. This order is being endangered again by hypothetical supersymmetric partners of all particles and murky dark matter and dark energy. Finally, string theory with its zillions of versions, perhaps realized in zillions of different worlds of the multiverse, kills the dream of finding a single unique principle governing the complex world. A multitude of outcomes renders this theory immune to empirical falsification, which, if we follow the definition by Karl Popper (1959), removes it from the realm of science, back into the ancient world of metaphysics and magic. Peter Woit (2006) had more to say on this in his book, whose title is taken from Pauli's quip: "Not Even Wrong".

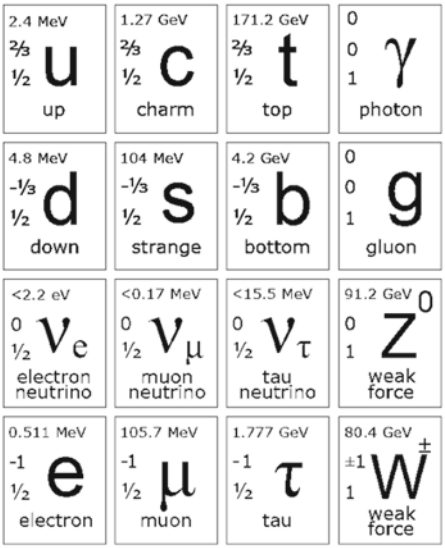


Fig. 1.7 Standard Model. *Three columns on the left*: generations of quarks (*upper two rows*) and leptons (*lower two rows*). *Right column*: bosons (gauge fields), with photons in the upper row

String theory virtually abolishes the very idea of elementarity, and spacetime along with it. So, in a different way, does another radical idea, reducing reality to the operation of cellular automata, a discrete computational tool initiated by John von Neumann (1951), which operates, not unlike old-fashioned computer games, by changing the states of cells (usually, in two dimensions) in a way that depends on the state of their neighbors. This can generate curiously evolving patterns, and can also be helpful in some model computations. The proponents of cellular automata, however, have gone much, much further. Konrad Zuse (1969) promulgated the idea of a “computable universe”, with the world being the result of a deterministic digital computation. Stephen Wolfram[3] in his voluminous book (Wolfram, 2002) hails cellular automata as a universal tool, and their study as a “new kind of science”. In the view of Gerard t Hooft,[4] reality is information, processed by a cellular automaton fabric operating at the Planck scale,[5] and the fundamental particles we know and love are emergent virtual particles (’t Hooft, 2016). In the poetic incantation of Ross Anderson (2016), the quantum vacuum is God’s computer, God’s bubble bath, or even God’s cryptographic keystream generator. Is this “replacing magic with mechanism” or going back to magic?

Time will tell whether all these great ideas are theories, fantasies, or speculations. The matter of the Standard Model is all we consist of. Moreover, we and everything around us consist of good old atoms. Nevertheless, we strive to grasp the ultimate foundations of being, be it God or a “God particle” or the ultimate structure of spacetime. Discreteness of spacetime, though formally an intrinsic feature of quantum mechanics, can be felt only on the minuscule Planck scale, which is many orders of magnitude removed, not only from our everyday experience, but also from the scales that any Earth-bound device might probe[6]. Hopefully, we are not living inside a computer game. Speculations about a program computing the Universe (or perhaps a universe) that should be as big as the Universe are just reiterations of an ironic fable by Jorge Borges about a map as big as the territory it covers.

1.4 Swings of Cosmology

In the early 1970s when I enjoyed publishing not-quite-Soviet papers in the popular science magazine “Znanie Sila” (Knowledge is Power) in Moscow, I composed a comical story about the creation of the World. The story was rejected as a parody on building socialism (which I didn’t mean at all) by the female commissar who had to keep the buoyant and unruly editorial staff in check. The text is lost, and I remember only the outline. The action took place in a Soviet-style research institute presided by a bland Soviet-style bureaucrat with a name and patronymic that sounded similar to the Russian rendering of the Tetragrammaton (not a four-letter word, which is a three-letter word in Russian). The staff had names hinting at pagan gods and were engaged in similar intrigues, while lazily planning worlds on the basis of old recipes, with the flat Earth resting upon something, etc. All went well till a junior research associate, called Pashka (an informal variant of Pavel, or Paul) Dyrkin (“dyra” is a hole, and you will guess who is intended by the phonetics), stole

---

[3] Formerly, the youngest MacArthur Fellow and the creator of the *Mathematica* software.

[4] His 1999 Nobel Prize in Physics was not given for these ideas.

[5] The unit of length, about 10^{-35} m, obtained by combining three fundamental constants, viz., the Planck constant, the gravitational constant, and the speed of light.

[6] John Wheeler (1998) used a money analogy to give a feeling of how huge this gap is: “a penny relative to the U.S. annual budget is a million times larger than the Planck length relative to the size of a proton”.

a *singularity* and ran away. As he was not qualified to deal with such an object[7], the singularity exploded, leaving in its wake the world as it was understood at the time of writing, with extreme contrasts of high densities and emptiness, heat and cold. Moreover, since the accounting was sloppy and goods were pilfered from the institution's storage, the ultimate fate of this world could not be determined (it was not known in the 1970s whether the Universe would expand forever or collapse in a Big Crunch).

Humankind developed a permanently expanding picture of the world, from naive myths to a permanently expanding Universe many billions of years old and many billions of light-years across, and even beyond that to a multiverse. The macroworld of cosmology, which, as we now believe, is deeply related to the microworld of elementary particles, has also passed through its cycles of simplicity and complexity. The neat Aristotelian system of seven planets rotating around the quiescent Earth was first supplemented by Ptolemean epicycles to better fit observational data. This was subsequently simplified by the heliocentric system of Copernicus (1543). The young Johannes Kepler (1596) came with a fantastic scheme for nesting the five Platonic solids, each encased in a sphere, to produce six layers that were supposed to correspond to the orbits of the six known planets (Fig. 1.8). He thought at the time that he had revealed God's geometrical plan for the universe. The perfect geometry of a circle or a sphere came, however, into contradiction with refined astronomical measurements by Tycho Brahe. To explain them, Kepler (1609) eventually came upon the idea of elliptical orbits with the Sun placed at one of the two foci. He created a formula in which a planet's rate of motion depended on its distance from

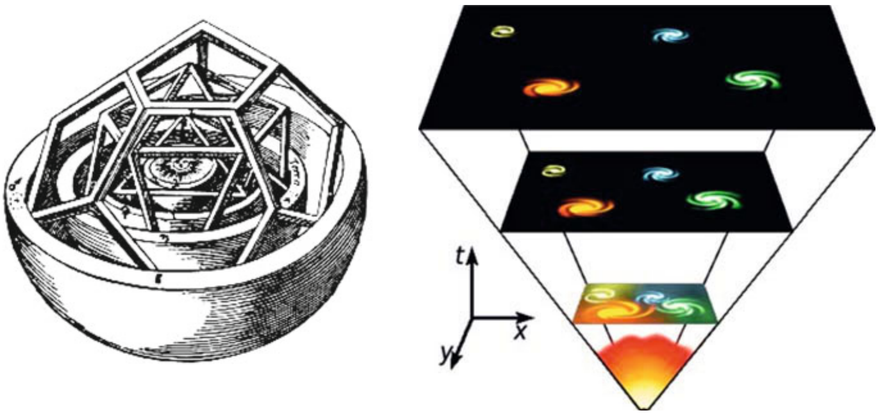


Fig. 1.8 *Left*: Kepler's nesting of the five Platonic solids. *Right*: Schematic rendering of the Universe expanding from the Big Bang. There is a hot opaque plasma in the lower corner, and spiral whirls represent galaxies

[7] Indeed, Paul Dirac was unable to overcome the singularities of relativistic quantum theory that he created. Richard Feynman, Julian Schwinger, and Shin'ichirō Tomonaga were awarded the 1965 Nobel Prize in Physics for the discovery of the renormalization method that solved this problem.

the Sun, which, as a symbol of God the Father, was the source of the motive force in the Solar System. This led to Kepler's law, stating that the square of the orbital period of a planet is proportional to the cube of the average distance from the Sun. The laws of planetary motion were soon rationally explained by Isaac Newton's theory of gravitation. Though Newton was no less mystically inclined, he concentrated his esoteric views on his passion for alchemy rather than on the mathematical and physical work we honor him for to this day.

Newton's theory reigned supreme even after the Sun was demoted to the status of an insignificant run-of-the-mill star, and the cosmological question shifted to the structure of the Universe as a whole. The first trouble surfaced in the experiment by Albert Michelson and Edward Morley (1887) proving that the speed of light – which was thought at the time to be transmitted through the aether filling up space, as sound is transmitted through air – remains invariant, independently of its direction of propagation relative to a moving observer. This made it necessary to consider time on the same footing as spatial coordinates. The paradox was resolved by Albert Einstein's special relativity theory – though Einstein claimed that he was not aware of these experiments and was driven only by mathematical beauty and logic – by the quest for simplicity (Einstein, 1905).

Since then, the cosmological question has turned into the question of the structure of *spacetime*. Still, no observations had contradicted Newton's gravitation theory when Einstein (1916) came along with his general relativity theory, which replaced Newton's gravitation with pure geometry. Einstein became world famous after his theory was confirmed by measuring small deviations in the orbit of Mercury and the bending of light rays by gravity during the solar eclipse of 29 May 1919. For him, this was rather a non-event. What significance could these measurements of tiny deviations have, compared to the great edifice of his theory? When asked what he would have done if the experiment had not confirmed his theory, Einstein ostensibly said: "*I would pity poor God*" – poor God indeed, who had not created the world according to this beautiful plan.

The Universe was now understood as spacetime governed by Einstein's relativity. Spacetime became malleable, a participant in the cosmic theater rather than just a scene where actors perform. This inspired my youthful verses (rhymed in the original Russian):

> [...] In those times,
> space was what the prophetic formulas are now dreaming:
> not just a page: a battlefield, a movie screen,
> the intersection of ages, embraces, and wounds.

Those times were our times, "when Icarus flew, made of wax, but with steel wings".

So far, Einstein's theory has stood as an unassailable fortress. There have been many attempts to modify it: adding a scalar field (Brans and Dicke, 1961), adding terms quadratic in curvature (Weyl, 1919; Vassiliev, 2005), or adding torsion, a non-propagating field that twists spacetime (Cartan, 1922; Hehl et al, 1976). Alternative theories that would predict gravitational waves propagating more slowly than light were disproved by the serendipitous coincident observation (Abbott et al, 2017) of

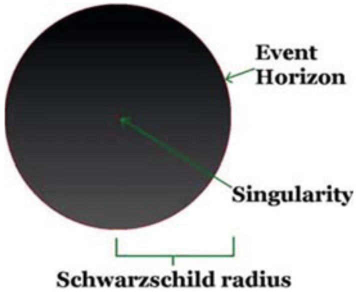


Fig. 1.9 *Left*: The Schwarzschild black hole. *Right*: Corona around a black hole (artist's concept by NASA/JPL-Caltech)

gravitational waves from a collision of two neutron stars and electromagnetic waves arriving almost simultaneously.

Solving equations of general relativity even in their standard form is a notoriously difficult task; usually, the form of a solution should be guessed and then justified by plugging this *ansatz* into the general equations. A very early solution was a *black hole* (Schwarzschild, 1916), a spherically symmetric spacetime metric around a mass so large that even light cannot escape from within its event horizon[8], located at the Schwarzschild radius (Fig. 1.9). The term itself became commonplace half a century later when it was popularized by John Wheeler[9]. Actually, a mass so large that light cannot escape it may already exist in the framework of Newtonian gravity! Pierre-Simon Laplace had almost arrived at a Newtonian black hole, perhaps hanging back because he thought this idea freakish. Einstein did not like black holes either, viewing them as a blemish on his theory. For a long time, their existence was doubted, and when the young Subrahmanyan Chandrasekhar[10] calculated that a star heavier than 1.4 solar mass should collapse into a black hole (Chandrasekhar, 1931), he was ridiculed.

More complex black hole solutions were obtained later (Chandrasekhar, 1983). Black holes formed from collapsed stars turned out to be common, and huge black holes with masses of many millions of suns may hide in the centers of the majority of galaxies, including our very own Milky Way. Black holes are far from being invisible, as they are surrounded by coronas radiated by accreting matter (Fig. 1.9). Even microscopic black holes, which might have been formed by fluctuations in the early Universe, have been imagined (Hawking, 1974), and theorists discuss the fate of information contained in matter falling into a black hole, and what happens to it when the black hole eventually "evaporates", which tiny things may do, while real astronomical objects of many solar masses would take forever.

[8] This "black shield" (if it may be so called) hides the naked singularity in the middle from the sight of an innocent observer, as required by the "cosmic censorship" hypothesis (Penrose, 1969).

[9] Wheeler (1998) recalled that Richard Feynman thought this term to be obscene, and wondered what associations Feynman made with it.

[10] He later won the 1983 Nobel Prize in Physics.

Alexander Friedman (1922) came up with a simple solution of Einstein's equation in the frozen Perm (Yuriatin in Pasternak's "Doctor Zhivago"). It described a homogeneous isotropic universe that expanded from a singularity, later called the Big Bang (Fig. 1.8). At first this was just a wild hypothesis, but then the visible world greatly expanded as Edwin Hubble, working with the world's largest telescope on Mount Wilson, identified hazy images called "nebulas" to be distant galaxies similar to our Milky Way. Measuring the dependence of the rate of expansion on the distance to these galaxies required two inputs. The velocity could be determined by the redshift of the light coming from distant stars, just as the sound of a fast car moving away has a slightly lower pitch. The distance could be measured with the help of "standard candles" – pulsating stars with a well-defined stable period and amplitude. Hubble (1929) found the rate of expansion to be proportional to the distance, and this dependence, extrapolated into the past, gave the age of the Universe. The coefficient of this linear dependence, called the Hubble constant, has since been corrected several times, although it remained constant up to observational errors until recently.

This did not yet validate the Big Bang hypothesis. The name itself was coined ironically by Fred Hoyle, who believed in an alternative scenario with continuous creation of matter eternally driving the expansion. The final proof was the discovery in the 1960s of the cosmic microwave background radiation, foretold by George Gamow, the drunkard genius, and detected by chance by Robert Wilson and Arno Penzias (1967), who were awarded the 1978 Nobel Prize in Physics. This radiation, coming (almost) uniformly from all directions, is the afterglow of the Big Bang: light strongly redshifted as its source recedes into the past. It comes from the earliest moment our optical tools are able to penetrate, the moment when the young Universe cooled off sufficiently for the opaque plasma to recombine into neutral atoms and the Universe became transparent. Later, more precise measurements discerned weak inhomogeneities in the radiation intensity caused by density fluctuations (Fig. 1.10).

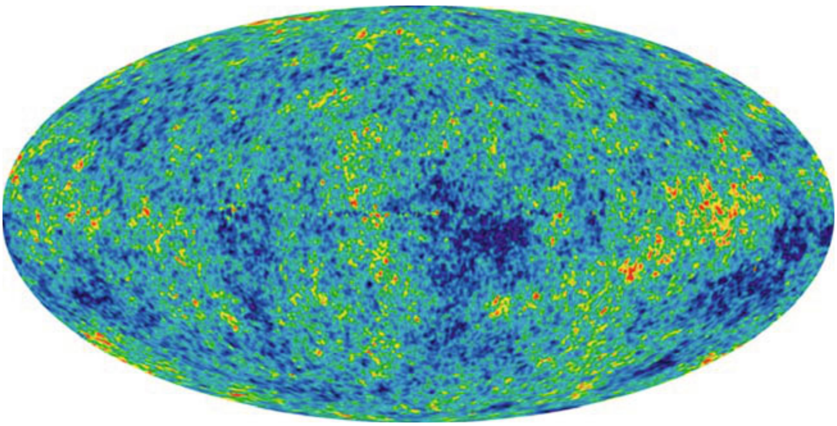


Fig. 1.10 Sky map of the intensity of the cosmic microwave background radiation (NASA image)

These would have developed into large density contrasts, leading eventually to the formation of galaxies and galaxy clusters.

Although no optical tools can penetrate closer to the Big Bang, whatever happened before that is relatively clear and rather bland, as there is fair evidence that the world was structureless and remained close to an equilibrium corresponding to temperatures gradually growing as one approaches the Big Bang. Mysteries start only when the Universe was so small that our knowledge of the physics of elementary particles is exhausted. Alan Guth (1981) came up with the idea that, in the very early stages, the Universe went through a period of exponential growth that had been caused by a phase transition from a "false" (metastable) vacuum to a "true" (stable) one. A rough analogy is rapid growth of an ice crystal from a nucleating seed in supercooled water – but here it is spacetime itself that grows, perhaps starting from a single Planck-size cell. Max Tegmark (2014) suggested a biomorphic metaphor: our baby universe grew literally like a baby human. "In this scenario", Tegmark writes, "our baby Universe grew very much the way you yourself did right after your conception: each of your cells doubled roughly daily, causing your total number of cells to increase day by day as 1, 2, 4, 8, 16, etc." What then had been going on before conception? What kind of parents did the baby have, if any? In the fourth century, Saint Augustin answered the question: Why was God idle for an eternity before creating the world? His answer complied perfectly with general relativity: before that, time did not exist.

The inflation scenario explained the homogeneity of the Universe, even among its furthest removed parts, which could have no mutual casual connection because they lie well beyond the visible horizon. The present and future state of the Universe as a whole appears now to be a simple story with a boring end, even though we are not familiar with all the characters and still less with the intricacies of their interactions. Dark matter and dark energy, both of unknown nature and apparently interacting with common matter only gravitationally (or, if in any other way, so weakly that it has not been detected so far), neatly balance one another in such a way that the Universe is (almost) flat. Dark energy drives the accelerating expansion of the Universe. This acceleration, previously unexpected, was proven by the efforts of two competing groups, who carefully measured the redshifts of supernovas in galaxies lying billions of light-years away (Perlmutter et al, 1999; Riess et al, 1998). This required a huge effort to calibrate these more powerful "standard candles" to establish the relation between the distance and luminosity of these huge faraway explosions, then painstakingly scanning the sky for these rare events. A typical undertaking of modern "Big Science", it involved work by dozens of junior researchers, fighting for funds and telescope usage, and tough intergroup competition, vividly described by Richard Panek (2011) – quite a contrast to the lonely contemplations of the great physicists of earlier times. The effort supplied reliable numbers but not insights. It came to a happy end, with the leaders of both groups awarded the 2011 Nobel Prize in Physics. Thankfully, Alfred Nobel had restricted the number of awardees to three.

The nature of the dark energy driving the expansion of the Universe is the major riddle. Einstein originally introduced a cosmological constant into his basic equation connecting the curvature of spacetime to the energy–momentum tensor. He tried thereby to make the Universe eternal. This was a mistake, since a static solution would be destroyed by instabilities anyway. However, the cosmological constant can be brought back to serve the same purpose of accelerating expansion. Another candidate for the dark energy is "quintessence" (the old aether again!). Unlike the cosmological constant, it would be a field changing in space and time. Current measurements fit the cosmological constant scenario rather better – but there is a big problem. Quantum mechanics predicts fluctuations of the vacuum that would generate a huge cosmological constant (Weinberg, 1989), about 120 orders of magnitude larger than the value compatible with astronomical observations. In any case, the fate of the Universe expanding at an ever accelerating rate is dull. As the distances to galaxies grow, they will recede beyond the cosmic horizon, and then the stars will recede as well:

> This is the way the world ends.
> Not with a bang but a whimper. (T. S. Eliot)

Our Sun will be long dead before this happens, and more massive stars will collapse into black holes which some believe to be wormholes to other universes. Perhaps the future will bring more surprises. The remote past, beyond the opaque curtain, is still a mystery, and complexity is bursting into the picture of a bubbling multiverse where anything is possible. The contradictory modern and postmodern attitudes can be seen in the two statements on the last blackboard of John Archibald Wheeler (Misner, Thorne, and Zurek, 2009): (6) "Omnibus ex nihil ducendis sufficit unum" (one principle suffices to obtain everything from nothing). (8) "Physics has to give up its impossible ideal of a proud unbending immutability and adopt the more modest mutability of its sister sciences, biology and geology". The first, is Einstein's uncompleted quest; the second, a retreat from his challenge to God. Other notes on the same blackboard underscore the humility of those who would stalk the secrets of creation: (5) "No explanation is an explanation that does not explain how the universe comes into being out of nothingness; not out of the vacuum of physics with its fluctuations and virtual particles, but out of nothingness. No laws, no particles, nothing". (16) "The laws of physics reveal as little about the deeper structure of the universe as the laws of elasticity reveal about the quantum mechanics of the solid state".

The single principle, if found (if it exists), would unify the structure of the Universe and the system of elementary particles. It might be hidden on the Planck scale, where it could be accessible to insight, although not direct experiment. This would still not necessarily mean that it was unverifiable and unfalsifiable, because its consequences might be tested, perhaps as weak effects, on scales available for observation. Einstein's general relativity might then be proven as an emergent theory rather than a fundamental one.

1.5 The Unreasonable Efficiency of Mathematics

Andrei Linde (2012), one of the leading proponents of cosmic inflation and multiple universes, starts his response to the Edge Foundation question about a favorite deep, elegant, or beautiful explanation by citing great men: "*The most incomprehensible thing about the world is that it is comprehensible* – this is one of the most famous quotes from Albert Einstein – *The fact that it is comprehensible is a miracle*. Paul Dirac said: *God used beautiful mathematics in creating the world*. Similarly, Eugene Wigner said that the *unreasonable efficiency of mathematics is a wonderful gift which we neither understand nor deserve*." He follows by contrasting the belief that "God created the universe and made it simple enough so that we can comprehend it" with the anthropic argument: "mathematicians and physicists can only live in those universes which are comprehensible and where the laws of mathematics are efficient".

Greek philosophers, the "mathematicians and physicists" of the day, and perhaps others before them who left us no trace, contemplated the same questions. For Plato, mathematics was the real world, and those ascending from the prisoners' cave went through the process of comprehension. "He will require to grow accustomed to the sight of the upper world. And first he will see the shadows best, next the reflections of men and other objects in the water, and then the objects themselves; then he will gaze upon the light of the moon and the stars and the spangled heaven; and he will see the sky and the stars by night better than the sun or the light of the sun by day" (Plato, *The Republic*, book VII). Those who learned about the real world "must be made to descend again among the prisoners in the den, and partake of their labors and honors", to be the guardians of the Republic.

Plato considered mathematics to be the most efficient and the most practical tool. But what was the mathematics that had to be learned? It was the numbers: "if Agamemnon could not count his fleet (and without number how could he?) he must have been a pretty sort of general indeed, [...] as I was saying, that arithmetic has a very great and elevating effect, compelling the soul to reason about abstract number, and rebelling against the introduction of visible or tangible objects into the argument. [...] The geometer is always talking of squaring, subtending, apposing, as if he had in view action; whereas knowledge is the real object of the study. It should elevate the soul, and create the mind of philosophy". Euclid's Geometry was, however, the most lasting (and most scientific) heirloom. Galileo (1623) wrote that "this grand book, the universe [...] is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures". Analytical geometry, bringing together geometric figures and numbers, was the major (if not the only) contribution René Descartes (1637) made to science. And numbers have reigned supreme since Newton's invention of differential calculus, reaching new heights through the work of the great French, German, and British mathematical physicists of the 18th and 19th centuries. Geometrical, or rather topological, methods have become prominent again today as a tool for studying the structure of both condensed matter and the Universe at large.

When it was discussed at Yale University whether to put more emphasis on languages or mathematics, Willard Gibbs, usually silent, said succinctly: "Mathematics is a language". This statement, perhaps fabled, has been repeated more than once, but it also has a less famous underside: mathematics is not science – it is a language. Freeman Dyson (1972) observed sadly that "the marriage between mathematics and physics, which was so enormously fruitful in past centuries, has recently ended in divorce". This could perhaps be better defined as a personality split: Euler, Laplace, and others were physicists as well as mathematicians, and were enriching the language they used. The 20th century mathematicians, with few notable exceptions, pursued formal linguistic problems unrelated to the real world. This gave them boundless freedom, but deprived them of the sense of direction distinguishing problems important for the world at large from problems important for mathematicians. The latter were famously set out by David Hilbert in 1900, and their solution has provided material for Fields Medals (the mathematical Nobel Prize, but much less richly remunerated) ever since. Hilbert, along with Gottlob Frege and Bertrand Russell, aspired to define mathematics logically; it was a bumpy way, encountering paradoxical things like "a set of sets which are not members of itself".[11] The incompleteness theorem of arithmetic proven by Kurt Gödel (1931) was hanging over these formal efforts like the sword of Damocles.

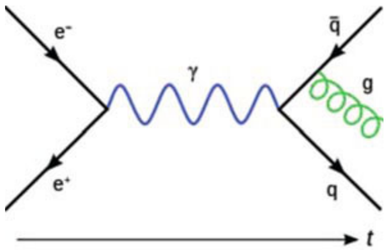


Fig. 1.11 In this Feynman diagram, an electron e^- and a positron e^+ annihilate, producing a photon γ that becomes a quark–antiquark pair, after which the antiquark radiates a gluon g

Physicists had little patience with this. Well before mathematicians had entered the fray, Newton wrote: "Mathematicians, who all discover, investigate, and prove, must be satisfied with the role of dry calculators and unskilled workers; the other, who cannot prove anything, grasps everything on the fly and has pretences to everything, carries all the glory of both his ancestors and his descendants". Richard Feynman, who was this kind of "other", quipped that "physics is to mathematics as sex is to masturbation". But new mathematical tools were needed for the new physics, and physicists had to deal with this themselves, as Newton did, inventing calculus when it was needed. Paul Dirac improvised the delta function, which is equal to zero everywhere except at a single point where it is equal to infinity. This function is very convenient for picking out a certain point within an integral. Mathematicians laughed at this, but then built up a whole theory of generalized functions to accommodate this freak and its relatives. Feynman was getting rid of infinities with the help of his diagrams (Fig. 1.11), which he himself at first considered childish, but which eventually became a deft tool that would be widely used by theorists – one can imagine what Hilbert would say about such doodles. The mathematical constructions of physicists became more and more sophisticated as the theories became

[11] Such paradoxes have a long history, cf. "a Cretan said that all Cretans are liers".

more abstract. String theories, divorced from experiments, became a kind of a mathematical exercise. They largely lack mathematical (as well as physical) rigor, but at least in one case a string theorist has won appreciation: Edward Witten was awarded the Fields Medal – but not the Nobel Prize, as the Swedish Academy always demands experimental evidence. Paraphrasing Feynman, we can say that multiverse theories are to realistic physics as a hook-up is to marriage.

1.6 Science, Revolutionary and Normal

In "The Structure of Scientific Revolutions", Thomas Kuhn (1962) famously distinguished between scientific revolutions leading to "paradigm shifts" and "normal" science operating within the confines of the leading paradigm and engaged in "puzzle-solving". The concept of paradigms has been exploited, against Kuhn's own convictions, by all kinds of postmodern relativists trying to reduce scientific facts and theories to the same non-decisive and temporary status as their own opinions and pronouncements, and even denounce them as tools of Western imperialists or Dead White Males. Setting these digressions aside, Kuhn's vision may be imaginative and poetic, but it does not truly reflect the way science operates.

The notion of paradigm shifts is not truly original, as it can be viewed as a version of the theory of punctuated equilibrium in biological evolution, evidenced by the sudden appearance of new species in the fossil records after prolonged periods of stasis (Mayr, 1954). Still earlier, Yury Tynianov (1924) discerned this pattern in the history of literature, contrasting it with the concept of gradual evolution prevailing in Darwin's theory as it was then understood: "not a systematic evolution but a jump, not development but displacement". In Tynianov's vision, new literary genres do not ripen within old ones, do not inherit from them, but displace them. Literature is like a colony of plankton floating on the surface of the warm sea of everyday life. A new form or genre is carried by deep sea currents; it is a feat of genius to sense this stream and bring it to the center of the colony. Once there, the new form spreads out and captures neighboring regions, but its potential is limited, as epigones follow in the footsteps of talent. Literary devices, once found, become automated and set in, while the older form is displaced to the periphery of the colony and drops out of the realm of literature, entering spoken language and everyday correspondence. Meanwhile, a new form emerges at the center of the colony. Thus, as the surface flux is centrifugal, it is closed by the undercurrent in the depths of everyday life. Further back in time, the early 19th century poet Yevgeny Baratynsky described the evolution of an idea from a poem – a young maiden, to a novel – an experienced woman, to journal polemics – an old impudent chatterbox.

Paul Dirac said: "In science one tries to tell people, in such a way as to be understood by everyone, something that no one ever knew before. But in the case of poetry, it's the exact opposite". Not really! More people understand Eliot's poetry than Dirac's equations, and human intellectual activities follow similar patterns. Physics is also a plankton colony, and the sea where it floats is nothing but Nature herself.

From the depths and into the center of the colony comes the stream originating scientific revolutions or paradigm changes. Just as literary forms are not compiled from the material of everyday life, physical theory does not arise out of raw experimental data: only the ordering skills of a genius can crystallize a new paradigm. By Kuhn's theory, a new paradigm, just like a new literary form, cannot ripen inside the old one; on the contrary, it arises (not without birth pains) when the prevailing conceptual framework is no longer able to account for newly discovered phenomena. This is what happened in the early modern period and again at the turn of the 20th century. The state of fundamental physics and cosmology in the early 21st century is different: there are more theories than solid facts.

Once the conceptual framework that stands in accord with the new facts has established itself in the center of the colony, it will begin to unfold. It will lead, on the one hand, to consolidation of the new system, with a more rigorous and harmonious formulation of its fundamental concepts, and, on the other hand, to its application to various complex phenomena. "Normal" science spreads out and grows incessantly as it elaborates upon its central idea. We can view the forthcoming Chaps. 2–4 as the unfolding of Newton's paradigm; we will talk about the birth and outreach of quantum mechanics in Chap. 6, while the cosmology discussed in Sect. 1.4 remains, before speculations congeal into a new paradigm, the unfolding of Einstein's theory. The tendency of "normal" science to proliferate directs it to ever more complex phenomena that may be insensitive to changes in the paradigm, as we will see in Chaps. 7–9. The periphery of science, where it merges into engineering, comes closest to practical applications. As journals are read by more people than poems, this area is most densely populated. It enriches everyday life and does not deserve the scorn of the melancholic poet. Here the analogy with Tynianov ends: the stream of science is not circular, but comes from Nature and flows into technology.

There is another sense in which the circularity of Tynianov's process does not apply to science: in a similar way to biological evolution, science possesses the arrow of irreversible advancement. A fitting linear direction can be found in the literary theory due to Northrop Frye (1957). The five literary modes following one another in the course of history are myth, legend, high mimesis, low mimesis, and irony. The protagonist of a myth is a god. The protagonist of a legend is a human acting under extraordinary circumstances and capable of performing miraculous deeds. The protagonist of high mimesis is of lower stature but still above the level of an ordinary human. The protagonist of low mimesis acts in an everyday environment, possessing all our weaknesses, passions, and hopes, and bound by the same social customs. Finally, the protagonist of ironic literature stands below our level by the heaviness of his situation, his moral qualities, the depth of his disillusionment, and the absurdity of his actions. Going down this scale, we pass from the eagle's eye view of the myth, indifferent to inconsequential detail, to the hazy vision of the legend, to high mimesis, which is like a commander observing the battlefield, to the impartial mirror of low mimesis, and to the magnifying glass of irony.

In a forgotten publication (Pismen, 1973a), I projected this scheme on science. The common source of literature and physics lies in creation myths. Thales and Pythagoras belong to this same mode, while the legendary stage spread from Plato through Aristotle to the medieval alchemists. Science proper began on the stage of high mimesis in the 16th to 18th centuries, while low mimesis in the 19th century brought detailed theories of fluid flow, elasticity, and transport processes, but also rational chemistry, thermodynamics, kinetic theory, and electromagnetism, and it continues to hold sway in hydrodynamics, material science, chemical theory, and related fields. The features of the ironic stage are seen, on the one hand, in esoteric theories of the highest and lowest reaches of physics and, on the other hand, in excessively detailed studies of earthly phenomena. In the remote reaches, string theories and multiverses bear distinctive features of the ironic stage; the appellation "ironic" is frequent in the ironic book on the end of science by Horgan (1996). Of course, categorizations or this kind are very loose; they are more metaphors than clear divisions.

Paradigm shifts are very rare. There have only been two scientific revolutions, hardly enough to build up a far-reaching theory, or perhaps only one: the emergence of quantum mechanics and relativity in the early 20th century. The first great breakthrough, by Galileo, Isaac Newton, Christiaan Huygens, and Robert Hooke, marked the *birth* of science rather than a scientific revolution. Steven Weinberg (2015) observed that "a modern scientist feels at home from 16th century onward but not before". Moreover, the word "scientist" did not even exist in the epoch of these great *natural philosophers*. The word was introduced by William Whewell (1834) in a book review. For the gentlemen natural philosophers of high mimesis, "the word scientist implied making a business of science; it degraded their labors of love to a drudgery for profits or salary" (Ross, 1962). Science proper did not exist before them, and after them it descended into a "drudgery" of low mimesis and beyond. Descartes deplored this descent, considering experiments to be "for the most part so complicated with unneeded details and superfluous ingredients that it would be very difficult for the investigator to discover their core of truth".

The vision of Plato or Democritus was poetic rather than scientific – not to mention Thales, who created a myth about water that differed little in its credibility from other creation myths. Aristoteles, the creator of a "paradigm" sustained longer than any other non-religious belief, would not care to carry out any experimental test of his statement that a stone thrown upwards keeps rising for a while, against its natural attraction to the center of the world, because it is supported by air, and would not even try to compare the amounts of water that could be contained in an empty vessel and in a vessel filled with sand. Technology was developed from the earliest times without any reference to philosophy. Perhaps Archimedes, with his famous "Eureka" to celebrate the results of an experiment in his bath, was the only scientist born before his time. If the Romans, instead of putting him to the sword, had set up a research institute for him, and later encouraged Hero of Alexandria to explore the capacities and applications of his steam engine, they would have been better prepared for the barbarian invasions, and we might even have enjoyed the Pax Romana to this very day. Instead, Europe's ascent, with some delay, was fueled by experiment and the inductive reasoning promoted by Francis Bacon (1620). As quipped by William Harvey, the discoverer of blood circulation, Bacon's was Lord Chancellor's science. It is also said that Bacon promulgated experiments but it was

Robert Boyle who did them, or rather Robert Hooke, his assistant, who was not a gentleman, did them for him. The famous "Knowledge is Power" is what all rulers value.

The birth of science was a singular event. It may not have been particularly decisive, in the sense that the absolute majority of people on this planet have no idea what the scientific method might be, and soccer players or pop singers are far more popular and influential than any scientist. Stephen Hawking was an exception confirming the rule, propelled to fame by his disability more than by his brilliance. He was a bestselling author, a media figure, even a character in the Simpsons. Eddie Redmayne won both the Golden Globe and Oscar for Best Actor for his brilliant portrayal of Hawking in the 2014 feature film *The Theory of Everything*.

But the birth of science was certainly decisive on the applied side: even the staunchest religious fundamentalists drive cars, launch rockets, fly in planes (and occasionally use them otherwise), and would not be seen dead without their smartphone. Only the Amish can be praised for being consistent in their practice and beliefs. This event was certainly decisive for science itself. The "change of paradigm" in the early 20th century did not abolish classical physics, which remains applicable and widely used within its well established confines. A great many technological applications (everything except smartphones in the above list) still operate within the classical paradigm, even though the media nowadays tend to apply the word "technology" to the likes of Google and Facebook and their lesser relatives, rather than to the more tangible technologies of old that we keep on using and improving.

All scientists through the ages, except for a few of the greatest who could be counted on the fingers of one hand, or let's say two and add toes to be generous, engaged in "normal" science in the context of a ruling paradigm – but the "puzzles" they were solving were both difficult and substantial. "Normal" science also oscillated between the simplicity of ignorance and the complexity of inscrutable phenomena, the simplicity of basic explanations and the complexity of detail.

Chapter 2
Continuum Mechanics

2.1 The Dry Water of the Age of Enlightenment

The basic tool of scientific theory was, and still is, Newton's differential calculus. Gone forever was the Greeks' inability to comprehend infinitesimals, Zeno's paradoxes with Achilles never catching up with a tortoise, and nobody ever attaining a fixed target, leading directly to the impossibility of motion. There is an infinite series: you first halve a distance, then halve the remaining half, and so on *ad infinitum*, never reaching your target. Everything becomes much clearer if you divide your path into infinitesimal intervals, which you may cross even with variable speed, and integrate your velocity as a function of time to obtain a definite and finite answer. A 21st century child, at ease with a speedometer in the family car and its relationship with the distance traveled, still battles with mathematical proofs of limit theorems, after moving on to Calculus 101, reliving the Ancient Greeks' troubles with understanding infinitesimals, and more often than not acquiring a lifelong aversion to math in the process.

Isaac Newton (1686) applied his calculus of "fluxions", now called time derivatives, to the motion of point bodies (including planets, on this level of abstraction). It was left to the 18th century mathematicians to extend the calculus to space and the three-dimensional material continuum. Bishop George Berkeley could not understand why an infinitesimal differed from zero and why, unlike zero, it could be divided by. Since he also denied the existence of any reality independent of consciousness, it would surely have been hard to persuade him. But Leonhard Euler, Daniel Bernoulli, Jean le Rond d'Alembert, and Joseph-Louis Lagrange did not hesitate to deal in infinitesimal fluid volumes and the infinitesimal forces exerted upon them by their equally tiny neighbors.

These were not the kind of mathematicians getting Fields Medals nowadays. They preferred computing to proving theorems, but this "sloppy" attitude (typical, in the modern view, of physicists rather than mathematicians) may have been what made possible Euler's lasting contributions to analysis, topology, and number theory, Bernoulli's analysis of probabilities presaging game theory, the creation of the

L. Pismen, *The Swings of Science*,
https://doi.org/10.1007/978-3-319-99777-3_2

calculus of variations by Euler and Lagrange, and the d'Alembert operator, born in music, finding a new life in special relativity which its creator could not have anticipated.

The newborn fluid dynamics was an elegant construction, and it could be viewed in two ways, used to this day in both theoretical and computational continuum mechanics: the Eulerian view, based on the balance of forces acting on an infinitesimal fluid volume located at a certain fixed point, and the Lagrangian view, following this element as it moves through space. In both formulations, the equations of fluid motion can be derived with the help of the calculus of variations by varying a certain energy integral with respect to the displacement of infinitesimal volumes. This method proved to be exceptionally fruitful, as it made it possible to derive equations describing the dynamics of more and more complex systems by changing the definition of energy and its mathematical expression. Notwithstanding the "change of paradigm", the equations of quantum mechanics and general relativity are derived using the same principle.

The legacy of Euler and Lagrange bears more influence on modern theoretical physics than on modern mathematics, but they were called mathematicians rather than physicists in their day. Eighteenth century fluid mechanics was far removed from the practical problems of the fledgling industrial revolution, even though the latter concerned with hydraulics and steam engines. Their equations lacked the most important component – viscosity! This was even a regression after Newton, who had carried out a series of experiments on bodies falling in water and established a linear dependence of the viscous force on the velocity – another law bearing his name, though now formulated in a different way, as a local relationship between the viscous stress and the velocity gradient. Was Euler aware of this? He left a comment indicating a zero net force on a body moving through a fluid at a constant speed – what came to be called the d'Alembert paradox – quite understandable in the absence of viscosity. Engineers would laugh at fluid mechanics of this kind, and Euler's advice on the water art constructions in Sanssouci Park did not contribute to the prestige of mathematics, although perhaps it was not his fault (Eckert, 2002).

This was the theory of the *ideal fluid* (Euler, 1757), called in jest "dry water". Indeed, such a fluid would not wet anything, as it would not stick to a wall. But the term "fluid" applies to air as well as to water, to gases as well as to liquids, and the viscosity of air can often be neglected. The Montgolfier brothers, unlike James Watt, wouldn't care about viscous effects. The ideal fluid is more than an abstraction: it is an example of a powerful scientific tool – *approximation*. Striving for absolute precision, for a proof without doubt, is sterile. Even sentencing in a court of law requires guilt to be proven only beyond a *reasonable* doubt. Physics goes still further: the result should be valid up to a reasonable approximation *under certain circumstances*, under relevant conditions. The approximation can be corrected when conditions change or when another answer is required. This does not mean that physics seeks a conditional, relative truth. It means only that it seeks truth at a relevant precision. The alternative is finding no truth at all, being stuck in the swamp of irrelevant complications.

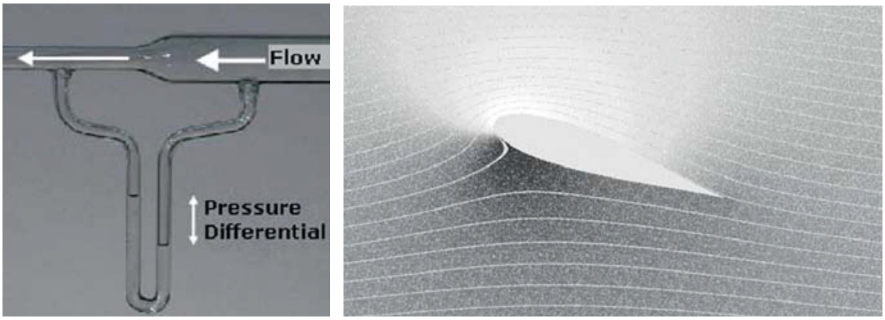


Fig. 2.1 *Left*: Decreasing pressure in a narrow section with a higher velocity. *Right*: Air flow around the Joukowski airfoil profile. Flow lines are shown in *white*, while *darker shades* correspond to a higher pressure

2.2 Hydrodynamics Airborne

One can hardly call viscosity an irrelevant complication in the context of 18th century technology. But there was at least one example of a theoretical prediction relevant for hydraulics: decreasing pressure in regions of higher velocity. This is Bernoulli's principle. It was rediscovered experimentally by Giovanni Battista Venturi as the "Venturi effect" (see Fig. 2.1), and later found application in a flow meter.

But more important practical consequences were to appear higher in the air. The circulation of air around a suitably designed airfoil causes the flow velocity under the airplane wing to be slower than above, whence the air pressure under the wing is higher, and this creates *lift*. When I was seven years old I sent a question to a radio show host: why can airplanes fly when they are so heavy? They answered that I would learn when I grew up. Indeed, it is hard to explain mathematical detail, but the basic principle is simple.

In the early 20th century, Nicolai Joukowsky (1910) computed a suitable airfoil profile bearing his name with the help of a conformal transformation in the complex plane. Omitting mathematical detail, conformal transformations can be illustrated by an age-old joke about a mathematician catching a lion in the desert. He (it's hard to imagine a female in this role) builds (or perhaps just draws) a cage and encloses himself therein. He then conformally transforms the desert plane, turning the interior of the cage into its exterior and *vice versa*. The result is that the lion (maybe not just one) will be in the cage while the mathematician roams free. In more formal terms, any analytical function of a complex variable can reshape the complex plane (spanned by the real and imaginary parts of a complex variable) in a certain way. The reader may not know what exactly an analytical function is, but what is important is that the equations describing an ideal fluid can be expressed in a complex form, whereas this becomes impossible once viscosity is added. Of course, viscosity is essential for computing *drag* upon a wing, and complicated computations taking into account detailed forms and even properties of materials are absolutely necessary

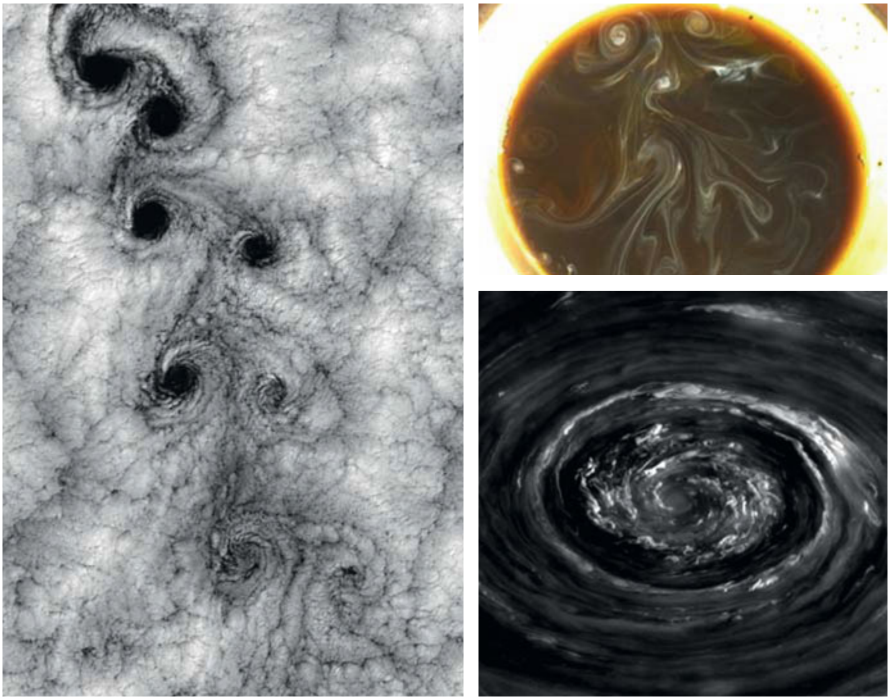


Fig. 2.2 *Left*: The von Kármán vortex street. *Right*: Milk vortices in a coffee cup (*above*) and the Saturn polar vortex (*below*)

to design an aircraft, while the principle of lift, following from 18th century theory, could actually be explained to a seven-year old kid.

Ideal flow is even more interesting in other respects. Hermann von Helmholtz (1858) found that a vortex sheet containing singular rotating flow patterns – *vortices* – is formed at a discontinuity of the flow tangential to a surface. These are easily observed as smoke rings or whirls in the wake of a boat or in the plughole of your sink. They are a major cause of drag on ships and aircraft. Hurricanes and tornados are giant vortices. Some examples are shown in Fig. 2.2.

Vortices can form vortex filaments, such as the vortex ring in the upper left panel of Fig. 2.3. A thin vortex filament can be represented by a line passing through its center, as in the lower panel. The shape of these vortex lines and their dynamics can be obtained analytically (Batchelor, 1967; Pismen, 1999). For example, the helical vortex in the central panel rotates around its axis. A funny example, known already to Helmholtz, is the motion of a pair of coaxial vortex rings. The self-induced velocity of a single ring decreases as its circumference grows. When they are paired, the trailing ring causes the leading ring to expand and slow down, while it itself contracts and accelerates under the action of the leading ring. As a result, the trailing ring passes ahead through the aperture of the larger ring, and starts to expand and

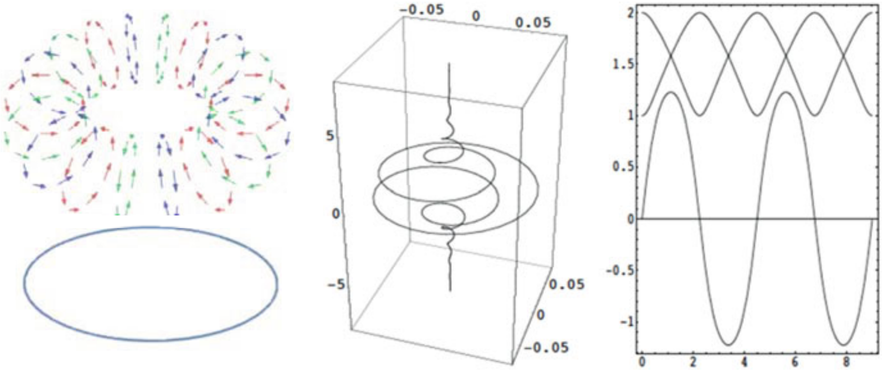


Fig. 2.3 *Left*: A vortex ring and its representation by a line. *Center*: A rotating helical vortex. *Right*: Radial oscillations (*upper curves*) and axial separation (*lower curve*) of leap-frogging vortex rings

slow down, as their roles interchange. The two rings continue leap-frogging as they propagate, as shown in the right panel of Fig. 2.3.

Such a solution certainly implies a lot of idealization. It presumes the vortex rings to be infinitesimally thin lines, and even in the framework of ideal hydrodynamics, it is destroyed by small deviations from the ideal circular shape. It also requires the fluid to be not only inviscid but incompressible, otherwise it eventually dissipates by radiating sound waves. Nevertheless, there was a lot of interest in solutions of this kind in the 20th century, especially following the discovery of superfluids, and later, Bose–Einstein condensates, an exotic state of matter observed at super-low temperatures when lots of boson particles occupy the lowest-energy quantum state (Sect. 6.6). Still more fancy constructions, extending to relativistic spacetime, are the hypothetical "cosmic strings" (Kibble, 1976). Further studies extended to Gargantuan computations of dynamic networks of vortex filaments, producing lovely pictures, but ignoring instabilities due to various imperfections and higher-order effects in earthly superfluids. Cosmic strings are still more vulnerable: should they exist at all, they will collapse, forming kinks moving with the speed of light, as shown in Fig. 2.4. Such a collapse would be spectacular, and astronomers will not fail to observe them.

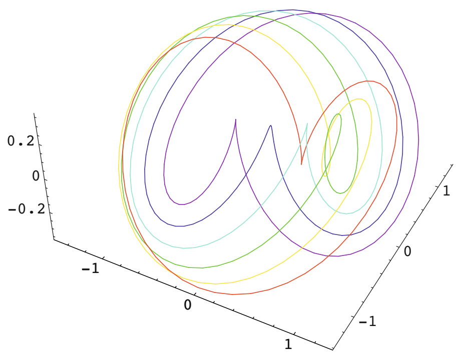


Fig. 2.4 Sequence of snapshots of a cosmic string. Half a time period is shown, starting from a cusped curve which evolves through a sequence of smooth curves into a symmetric curve with the cusp turned over

On a more fundamental level, William Thomson (1867), the future Lord Kelvin, imagined the atom being a vortex in the aether. The idea was popular at the time, before aether was abolished by Einstein, but it may still be resurrected at a deeper level. In the modern view, the *vacuum* is not just empty space, but a medium bristling with virtual particles that can be viewed as a more sophisticated version of the 19th century aether. It may even undergo phase transitions to a vacuum of another kind, and there were rather exorbitant fears that CERN's Large Hadron Collider (LHC) might trigger nucleation of an alternative vacuum that would spread out and annihilate our world. LHC is certainly not that powerful. In 2012, it hardly managed to generate events interpreted as a trace of the Higgs boson – the "God particle" that many physicists hoped would not be seen, to make possible a more imaginative closure of the Standard Model – see Sect. 1.3. Roger Penrose (1960) proposed the concept of a *twistor*, which describes the kinematical structure of a spinning massless particle, as the basic element of spacetime, with the "twistor space" replacing conventional geometry. Some physicists believe that, at distances approaching the minuscule Planck scale of quantum gravity, quarks and leptons will be revealed as topological singularities of the vacuum – Kelvin's idea extended into the depths of spacetime and beyond.

René Descartes imagined vortices (set in motion by God) filling up space and carrying the Sun, the stars, and the planets in their orbits. This theory (if it can be called so, as it had no rational support) precluded, and was made obsolete by, Newton's theory of gravitation. Descartes' speculations nevertheless extended far into the future. Could he have had in mind all these developments when he modestly observed: "I hope that posterity will judge me kindly, not only as to the things which I have explained, but also to those which I have intentionally omitted so as to leave to others the pleasure of discovery"?

2.3 Viscosity and Boundary Layers

Viscosity was brought into hydrodynamic theory in the early 19th century by Claude-Louis Navier (1823). The name of George Stokes (1845), who provided a more rigorous derivation, was later adjoined, to label the well known Navier–Stokes equation. Its restricted version is the Stokes equation, which retains only viscosity but not inertia, and is therefore the direct opposite of the Euler equation. The transition from Stokes to Euler through Navier takes place as the dimensionless Reynolds number, proportional to the velocity and a characteristic length and inversely proportional to the viscosity, increases from zero to infinity.

Helped along by Stokes's contemporaries William Thomson (Lord Kelvin), Lord Rayleigh, Osborne Reynolds, William Froude, George Biddell Airy, and Horace Lamb, Britain ruled 19th century hydrodynamics just as she ruled the waves[1], and

[1] This does not imply that other powers did not navigate, and does not diminish the accomplishments of Siméon Poisson, Adhémar de Saint-Venant, Hermann von Helmholtz, and Joseph Boussinesq.

in fluid dynamics, the contributions of Geoffrey Taylor, George Batchelor, James Lighthill, Keith Moffatt, and others, helped to keep 20th century British mathematics applied, while French mathematics was driven into a pure dry landscape by "General" Nicolas Bourbaki[2].

The Stokes equation could handle slow (creeping) flow rather well, and Stokes himself used it to obtain important results, such as computing the velocity of a particle falling through a liquid. In our own time, the Stokes equation is the main tool for studying the motion of micro-organisms and tiny artificial swimmers that will perhaps soon be able to deliver drugs to programmed locations by travelling along blood vessels. This equation does not contain the time derivative, and therefore, although it describes a dissipative process, is fully reversible. If, say, you put a tainted blob into a viscous liquid between two cylinders and rotate the inner cylinder gently, the blob will be stretched and thinned around the circumference, but when the direction of rotation is reversed, it will collect itself back into its original shape. Only very slow diffusion of the paint will blur the blob's surface. Contrariwise, flow obeying the Euler equation, although it originates in reversible Newtonian dynamics, is *irreversible*. Fast flow has too many degrees of freedom and tends to be chaotic, and one cannot reverse a jet breaking into a myriad of streaks and blobs, just as one cannot put Humpty Dumpty back together again.

Hydrodynamics is notoriously difficult for analytical studies[3]. On the one hand, this has prompted exquisite mathematical work, while on the other, it has necessitated difficult computations in practical applications. John Bernal (1954), an ardent communist but outstanding crystallographer, wrote at the dawn of the computer age that computational devices were "certainly to alter radically the whole of our thinking about quantitative methods in calculations in the same way as, and to a far greater extent than, the adoption of Arabic numerals did in the late Middle Ages". Computational fluid dynamics has even penetrated the entertainment industry, as creators of video games use the Navier–Stokes equation to make the artificial environment seem more natural. In the words of Roger Penrose (2013): "Computer simulations can lead to hugely impressive imitations of reality, and the resulting visual representations may be almost indistinguishable from the real thing, a fact that is frequently made use of in realistic special effects in films." Imitating visual effects is, however, much easier than computing essential characteristics, such as viscous drag or intensity of heat and mass transfer in real flows that are elaborately structured on short scales. The difficulties caused by the hierarchy of time and length scales grow and multiply in attempts to simulate very large systems, such as atmospheric phenomena, ocean currents, or the structure of galaxies. The climate models so boldly presented by climate scientists, in a bid to justify their grants, show a wide spread of predictions. Thankfully, as in the tale of Nasreddin promising the Emir to teach a donkey to speak in 20 years, checks on the results are safely postponed. Weather can be reliably predicted only for a few days and only where and when it is

[2] The collective *nom de plume* of a prominent group of French mathematicians.

[3] A thorough historical account of the development of hydrodynamics, including basic formulas and an extensive bibliography, has been compiled by Olivier Darrigol (2005).

stable (more on this in Sect. 9.1), and this can be only marginally improved on with the help of growing data caches and computing power.

Computer pioneers, Alan Turing and Konrad Zuse, working on the two sides of the bloodiest conflict ever, might have taken computer capabilities too seriously, as some their followers still do. I mentioned in Sect. 1.3 a megalomanic dream of a "computable universe" – but we don't need to rise to the skies or dive into the CERN tunnel to appreciate the interplay of widely separated scales. The flow of humble water along a plane can teach this lesson. Stokes prevails near the wall where the fluid is slowed down by strong friction, while Euler reigns far out, and the two limits need to be reconciled. Of course, one can solve the Navier–Stokes equation numerically, as now proposed by commercially available program packages, like the Fluent computational fluid dynamics (CFD) software. But such packages are problematic in many ways. First of all, any computation can only be carried out in a finite domain, although this doesn't trouble those computing universes. Sharp gradients that occur near the walls can be resolved by a denser mesh, but this is hard to do when they emerge in the bulk as well, and the treatment of sharp moving interfaces, such as Helmholtz's vortex sheets, is always inadequate. Turbulence can be accounted for only with the help of simplified models. More sophisticated but user-unfriendly programs are written by graduate students and postdocs, and the results of their computations, rarely verified, find their way into scientific journals.

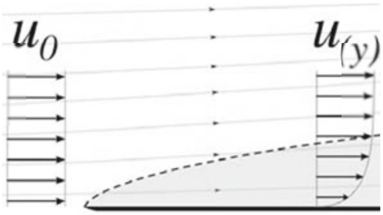


Fig. 2.5 Velocity profile in a laminar boundary layer

Neither software nor hardware were available when Ludwig Prandtl (1904) came out with the boundary layer theory. The basic idea, anticipated in the late 19th century by William Rankine, William Froude, and even the chemist Dmitry Mendeleev (Tani, 1969), and also applicable to diffusion, heat transfer, and the distribution of the electric potential (see Chap. 3), but never formulated mathematically before Prandtl, is simple. One needs to account for viscosity (or diffusivity or an electrostatic force exerted by the wall in other problems) only close to a boundary, and only the gradient in the normal direction is important (Fig. 2.5). Far from the wall, ideal flow equations can be used. The problematic part is matching the "inner" and "outer" regions. Elaborate analytical matching methods applicable to a variety of problems were developed by 20th cen-

Fig. 2.6 Jackson Pollock at work (photo by Hans Namuth)

tury applied mathematicians and physicists, but this skill appears to be withering away as younger generations become more and more infatuated by numerics. Of course, analytical methods are severely limited. Even computing an inviscid flow, or bulk solutions in transport problems, is possible only in simple geometries, and only when the flow is laminar rather than turbulent. Asymptotic matching is more an art than an algorithmic technique and is never guaranteed to work. I wish machines could be trained in this skill.

Sydney Goldstein (1969) writes: “Prandtl once told me that he had considered building an analogue machine, but came to the conclusion that water itself was the best. It is reported that [John von] Neumann once said that there would be no further need for experiment – high-speed computation could take over. I suppose both were incorrect; we need both experiment and computation”. Every September, the Journal of Fluid Mechanics publishes the Gallery of Fluid Motion containing dazzling pictures, both experimental and simulation snapshots, and both exceeding in their sophistication and aesthetic appeal anything that abstract expressionists could offer. The technique of Jackson Pollock was essentially fluid-mechanical (Fig. 2.6), and the randomness of his splashes lay at the core of his appeal. Alas, analytical theory cannot rise to these artistic heights.

2.4 Elasticity

There are similarities in setting the balance of forces on infinitesimal elements of solids or of fluids. Quite naturally, Euler, Daniel Bernoulli and his father Johann, and also Navier were engaged in both problems. Of course, fluids and solids behave in completely different ways: in fluids, the stress is caused by the gradient of velocity, and follows Newton’s friction law, while in solids the origin of stress is strain, or the gradient of displacement, and their linear connection is determined by the law put out by Hooke (1678). Both these laws originated in experiments, the former on solids moving in water, and the latter on the response of strings to an applied force, and they still had to be extended to the mechanics of continua.

It was already the early 19th century when Augustin-Louis Cauchy (1828) defined the notion of *stress*, and Navier (1827) formulated the equations of elasticity in the differential form, with Hooke’s law converted into a linear relation between stress and strain. The analytical linear elasticity theory was built up by British, French, and German mathematicians during the 19th century, as summarized in the classical treatise by Augustus Love (1892)[4]. A parallel development was the introduction of special functions bearing the names of Bessel, Lagrange, Legendre, Airy, and others, which were an indispensable computational tool at the time, and hopefully will not be forgotten, as modern symbolic computation software knows how to tackle them, though sometimes mixing up their branches in the complex plane. In some ways, elasticity theory is simpler than fluid mechanics. In fluids, lacking

[4] For a playful mathematician’s pun, see “The Ambiguous Twist of Love”, a paper by Alexander and Antman (1982).

cohesion, the strain is unlimited, and the theory is essentially dynamic. In solids, the strain is typically small, and the classical theory is linear, and it is not dynamic, concentrating on equilibrium configurations under the action of applied forces.

Practical needs called for a description of the way slender bodies – beams – would deform, rather than computing the entire distribution of the deformation field in a three-dimensional volume, which would have been quite impossible to achieve at the time. Abbé Edme Mariotte observed that the flexure of a beam involves the extension of half of its fibers and compression of the other half, these together determining its resistance to bending. The beam problem had already been posed by Galileo, and later developed mathematically by Euler, Daniel Bernoulli, and Charles-Augustin Coulomb. The latter was also the first to consider the resistance of thin fibers to torsion when he was inventing his torsion balance (Coulomb, 1784). The theory of beams was instrumental in constructing splendid 19th century bridges and fledgling skyscrapers.

Fig. 2.7 Vibration of the Tacoma Narrows Bridge prior to the 1940 collapse

Bridges are subject to vibrations that can have dire consequences, as in the collapse of suspension bridges in Broughton in 1827 and Angers in 1850, due to resonance with marching soldiers, reiterated less gruesomely on the London Millennium Wobbly. Vibration due to high winds caused the famous collapses of the Tay Bridge in 1879 and the Tacoma Narrows Bridge in 1940 (Fig. 2.7). Alfred Clebsch (1864) formulated the general theory of free vibrations of solids, bringing together and improving upon earlier results for vibrations in strings, plates, and membranes due to Poisson, d'Alembert, and others. The theory could help to understand this kind of disaster but, as we know, it did not prevent them: engineering and cutting edge mathematics rarely come together.

More elaborate late 19th century theories anticipated new challenges in the next century. Natural rubber was already introduced to Europe in the 18th century, but it was as yet an insignificant exotic material. Nevertheless, some of its properties, quite different from metal and wood, could have attracted attention. First comes the possibility of large deformations; second, a nonlinear dependence of stress on strain; and third, and most importantly, creeping under continually applied force. Rubber is not alone in the class of *viscoelastic* materials that combine, in differing degrees, the properties of solids and liquids. Glass gradually transforms from an elastic solid to a viscous fluid passing through a viscoelastic phase as it is heated, unlike crystalline materials, which undergo a sharp transition at a melting point. All kinds of ointments which would now be called colloidal materials are not quite solid and not quite liquid, as we can feel when squeezing toothpaste from a tube.

The investigation of the properties of such materials, now referred to as *soft matter*, has become one of the liveliest branches of applied physics in the 21st century, in particular, in connection with biophysical applications (more on this in Sect. 9.2). Much earlier, they did not escape the attention of James Clerk Maxwell who, though dying at the age of 48, before what had been called the *acme* (apex, or prime of life) by the Greek philosophers, deserved to be called the greatest physicist of his century. The Maxwell model, remaining the most widely used model of viscoelasticity to this day, relates the rate of change of strain with time to both the stress and its rate of change with time. A Maxwell material behaves as an elastic solid at short times, but flows as a liquid at longer times. Another celebrated scientist, William Thomson, came up with an alternative, known as the Kelvin–Voigt model, relating stress to strain and its rate of change with time. This reversal of the positions of strain and stress in the mathematical formula causes the material to be essentially solid, as it bounces back to the initial state after stress is removed, though this does not happen immediately, as it would in an ideal elastic material. A plethora of more complicated models, containing more parameters than just the elastic modulus and viscosity, were suggested later in attempts to describe the behavior of soft materials with a greater, but never attained, precision.

The study of large deformations prompted a modified definition of stress by Gabrio Piola and Gustav Kirchhoff, relating it to a reference configuration rather than to the current one; it coincided with the Cauchy stress only when deformations were small. The mathematical theory of large deformations was further developed through the 20th century, becoming overgrown with ever more detail, such as accounting for anisotropies in materials. Both the strain and stress became *tensors* of second rank (square arrays of coordinate-dependent numbers), or more precisely, tensor fields. The proportionality constant in Hooke's law, first split into the shear and bulk moduli, which related to volume-conserving and compressive deformations, respectively, was fattened up in anisotropic bodies to a fourth rank tensor containing in the general case, when we take symmetries into account, 21 parameters – too clumsy to be used in most situations.

Even when local deformations are small, large deformations readily occur in slender elastic bodies, such as almost two-dimensional thin sheets and shells or almost one-dimensional rods and filaments. Sheets, when deformed, are compressed on the concave and stretched on the convex side, and these deformations can be integrated to determine their flexural rigidity and reduce the three-dimensional elastic equations to the two-dimensional Föppl–von Kármán equations. The two-dimensional geometry of shells is the simplest non-trivial realization of the Riemannian geometry that later provided the mathematical basis for Einstein's general relativity, and a shortest path (geodesic) on a shell is determined by the same equations as a light ray in curved spacetime. The Föpple–von Kármán equations provide an example of *dimensional reduction* – from three dimensions to two. The various theories of higher-dimensional worlds, from the five dimensions of Theodor Kaluza and Oskar Klein in the early 20th century to the 10 or 11 dimensions needed to host fancy superstrings, should similarly eliminate extra dimensions to project their

Fig. 2.8 *Left*: An image of a naturally twisted ring filament. *Right*: A screenshot from the trailer for "Ratatouille", featuring computer-simulated hair

wonders onto four-dimensional spacetime – tasks never properly accomplished so far.

Continued modern interest in slender elastic bodies is driven by soft matter applications. In nature, trees, shrubs, and herbs grow undulating leaves and filaments curling in manifold ways which do not escape the attention of biophysicists. The growth itself is the cause of stresses determining the variety of shapes. Deeper down, nature manages to pack the meter-long double spiral of DNA into a tiny space, and to carry this out in a way allowing the chromosome to uncover particular genes upon request when they have to be expressed – still a physical and mathematical riddle.

On another level, the cosmetic giant L'Oréal was sufficiently interested in the elastic properties of hair to hire a young physicist as a consultant, and the same physicist was instrumental in writing a program to create the smart and agile rat in the animation *Ratatouille*[5] (Fig. 2.8). In his novel *Generation P* published in Russian in 1999[6], the Russian writer Victor Pelevin imagined programmers creating politicians in this way, less lovely than this particular rat, but also supposed to exist only on screen, while more powerful forces govern the world. The protagonists of this novel envied their American counterparts having more powerful computers that allowed them to render the US President's *chevelure* in a more realistic way.

[5] It was Basile Audoly, also distinguished by the Ig Nobel Prize, awarded for his imaginative work on breaking dry spaghetti (Audoly and Neukirch, 2005), and by the most advanced modern book on elasticity coauthored with his former mentor (Audoly and Pomeau, 2010).

[6] English translation by Andrew Bromfield, entitled *Homo Zapiens*, Viking Penguin (Pelevin, 2002).

2.5 Waves

Vibrations propagate through elastic bodies, structures vibrate, and strings vibrate, producing sound waves that propagate through air and then through the cochlea in our inner ear, where our brain converts them to music; winds drive sea waves that break in shallow water, and the trembling Earth generates tsunamis. In Sect. 2.2, I briefly mentioned sound waves carrying away the energy and momentum of vortex rings. Sound is a longitudinal compression–rarefaction wave. It requires the medium to be compressible but, perhaps counter-intuitively, the easier it is to compress, i.e., the more sensitive the dependence of density on pressure, the slower the speed of sound. It propagates more slowly in air than in water, and in an incompressible fluid, the speed of sound is formally infinite. The changes in pressure and density are small under typical conditions, whence sound waves can be described by linear equations. Sometimes, however, the conditions are not typical: an airplane approaching the speed of sound creates a very strong perturbation – see Sect. 3.6.

The kind of wave problem that could be treated using the tools of the theory of ideal incompressible fluids was propagation of waves on the surface of water. Even then, further conditions had to be imposed by Joseph-Louis Lagrange (1788) to arrive at analytical results: the amplitude of the waves had to be small, to keep the problem linear, and the water layer had to be shallow. Lagrange justified the latter assumption, not quite correctly, by observing that the motion is largely confined to the layers near the surface. The shallow water approximation assumes that pressure is almost constant over the depth of a layer and the velocity in the vertical direction is much smaller than along the surface. This is another kind of dimensional reduction, like the one we saw in thin shells; it had already been used by Daniel Bernoulli (1738), and Pierre-Simon Laplace (1798) applied it to the theory of tides, which also took into account friction. It is widely used to this day (Oron, Bankoff, and Davis, 1997), often providing the only way to treat the problem rationally. The two assumptions may, however, turn out to be incompatible: the amplitude of the wave should be small *compared* to the depth; we see what happens when this condition is violated whenever we watch waves breaking on a beach – more on this in Sect. 3.6.

Kozma Prutkov, the creation of three hilarious cousins in mid-19th century Russia (cf. General Bourbaki), advised: “Throwing pebbles into the water, look at the ripples they form. Otherwise it will be an empty occupation” – but scientists not only observe, they also compute the ripples! Laplace (1776), one of the greatest mathematicians ever[7], posed exactly the same problem: what is the motion following a localized initial disturbance of a liquid surface? What we usually see are gravity waves, driven by gravity trying to restore the equilibrium shape of the water surface when it is displaced by a pebble splash or wind shear; similar waves can emerge in the depths of the sea, on the boundary between layers of different density. Another kind are *capillary* waves, driven by surface tension that tends to flatten the surface by capillary pressure pushing down convex and lifting up concave patches of the surface (more on this in Sect. 5.3); such waves dominate at shorter wavelengths.

[7] Uniquely, both Napoleon and the restored Bourbon king adorned him with nobility titles.

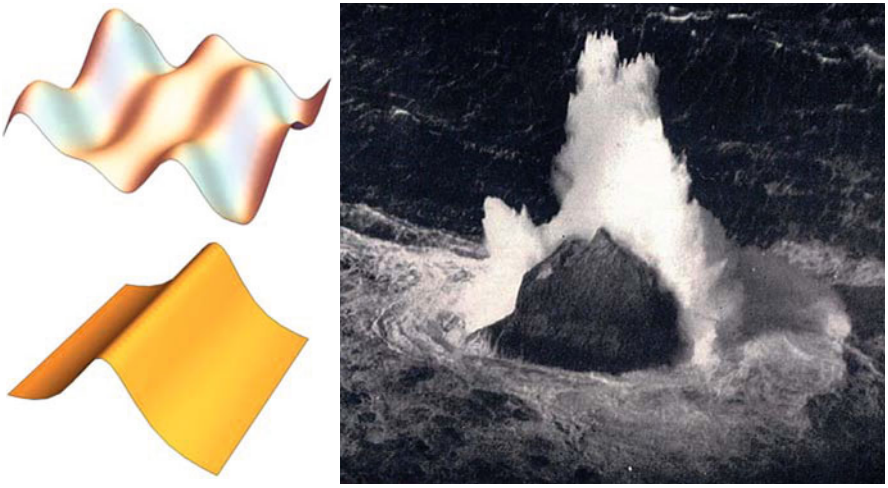


Fig. 2.9 *Upper left*: A surface wave pattern. *Lower left*: A solitary wave. *Right*: A 1943 photograph of a 50 m wave breaking over the islet of Rockall

Original theories of water waves were clumsy, as the main mathematical tool for the description of linear waves had not been yet discovered. This was Fourier analysis, which represents functions of spatial coordinates and time as sums of elementary waves with certain frequencies and wavelengths, expressed in the simplest case by trigonometric functions[8] or as an integral over a continuum of frequencies and wavelengths. As such elementary waves, or different Fourier components, add up and interfere, they can combine to form a variety of shapes for the water surface. While individual Fourier components propagate with their own speed equal to the frequency multiplied by the wavelength, they propagate together with a common velocity called the group velocity. Of course, what propagates is a wave of the disturbance, while individual water parcels, like air in a sound wave, just move back and forth.

The waves may combine into a *solitary* wave, a single bump rather than the wave train we usually observe. Such a wave was observed by John Scott Russell (1834), a young engineer riding along a narrow canal near Glasgow. Such waves, called *solitons*, later became the subject of intense research using model equations: the Korteweg–de Vries equation, an idealized equation for shallow water waves first derived by Joseph Boussinesq (1877), and the nonlinear Schrödinger, or Gross–Pitaevskii, equation describing ideal fluids and Bose–Einstein condensates. A magical property of solitons was only discovered when the first computers became available. In 1955, Enrico Fermi, John Pasta, and Stanislaw Ulam, modeling a chain of frictionless springs on the MANIAC I in Los Alamos, observed an unexpected recurrence phenomenon: the system periodically returned to its initial state. A decade

[8] Special functions may be used instead, in suitable geometries.

later, Norman Zabusky and Martin Kruskal (1965) modeled a continuous version of the same system, which they soon learned was nothing but the Korteweg–de Vries equation, and saw that when solitons collided, they passed through each other, emerging almost wholly unscathed from these encounters.

Of course, this magical property is highly fragile: not only must all dissipation be absent, but the equations must be *integrable*. Nevertheless, soliton solutions weakly perturbed by adding small realistic corrections are useful for understanding nonlinear phenomena that could otherwise only be modeled numerically. One of the mysterious phenomena that may be related to solitons are *rogue* (called also freak or monster) waves, observed occasionally on high seas and even on lakes, sometimes even in calm weather. Such waves, which can be up to 30 meters high (see Fig. 2.9), have wrecked cruise ships more than once, as recorded between 1901 and 2014; earlier and smaller ships may have been sunk without leaving any record. These waves are commonly modeled by the same soliton equations, also applicable to optical guides and lasers. Their theory was the topic of a Discussion & Debate issue (Akhmediev and Pelinovsky, 2010) that even contained a paper "Could rogue waves be used as efficient weapons against enemy ships?". Related applications can be still more whimsical. John Wheeler, fascinated in the 1950s by his beautiful but unrealizable idea of "geons", waves contorting spacetime into particles by their energy, imagined particles as embodiments of fantastically strong gravitational rogue waves (Wheeler, 1998).

The most important of all waves is *light*. The wave theory of light was originated by Christiaan Huygens (1690). The alternative was the corpuscular theory due to Newton (1704). The two theories were much later unified by Albert Einstein (1905a) in one of his *Annus Mirabilis* papers[9]. That paper on the photoelectric effect antedated quantum mechanics, which would subsequently also declare electrons to be both particles and waves (Sect. 6.2). For some time, neither the wave nor the corpuscular theory was preferred, as both only predicted propagation of light along straight lines with a finite, though apparently very high, speed. Huygens observed birefringence, occurring in anisotropic materials and depending on the polarization of light, as was explained much later by Augustin-Jean Fresnel. But Huygens thought light to be a longitudinal wave, similar to sound, whereas only *transverse* waves, perpendicular to the direction of propagation, can be polarized. The dispute was decided in favor of the wave theory by the double-slit interference experiment of Thomas Young (1802). But how light waves could be transverse remained a mystery.

Unlike fluids, solids can transmit, alongside sound, transverse waves of alternating shear stress. Such waves were first studied by Siméon Poisson (1816). James Clerk Maxwell (1873), unifying electricity and magnetism, showed that light waves were transverse, as were all electromagnetic waves, like the radio waves soon to be generated by Heinrich Hertz (1893). But what do such waves propagate in? Up to then, propagation of waves had required a medium. The sound made by Neil Armstrong's first steps on the Moon could not be heard, and water waves need a water

[9] It was characteristic of the Nobel Committee's cautious attitude to theories that he was awarded the Nobel Prize in 1921 for this theory explaining the photoelectric effect, rather than for general relativity, although the latter had already made him famous by then.

surface. The hypothetic medium in which electromagnetic waves were supposed to propagate was the luminiferous aether. Since electromagnetic waves are transverse, the aether had to be solid. George Stokes (1880), who contributed not only to fluid dynamics but also to theory of diffraction of polarized light, took this very seriously. The problem was to understand how the planets could move through a solid medium. Stokes suggested that the aether would be like a very dilute jelly to allow the planets to travel freely through it. A modern expert in material science would have noticed a contradiction since, on the one hand, a dilute jelly is too feeble to sustain vibrations, and on the other, even a tiny resistance would slow down the planetary motions over astronomical times, enough to cause the planets to fall into the Sun. Kelvin's theory of aethereal vortices did not fit either, as it assumed the aether to be a fluid. Would it be viscoelastic as earthly gels are?

As if this were not enough, further trouble came with the experiment by Albert Michelson and Edward Morley (1887), proving that the speed of light remains invariant, independently of the direction of its propagation relative to the moving observer. This contradicted the idea of a preferred coordinate system based on aether, and made it necessary to consider time on the same footing as spatial coordinates. Einstein's special relativity theory killed the idea of aether forever (see Sect. 1.4). With the aether abandoned, Maxwell's equations not only survived, but became even more elegant and compact in their four-dimensional formulation. Here, the electric and magnetic fields in vacuum are brought together as the six independent elements in the antisymmetric electromagnetic field tensor, which is related by a single equation to the four-dimensional current vector. I saw the four Maxwell equations, shown at his monument in Edinburgh (Fig. 2.10), nailed to an office door in the Physics Department as God's pronouncement on the second day of creation: "And God said", the equations follow, "and there was light". There are also T-shirts with this design. The Almighty could have been more succinct, expressing it in a single short line.

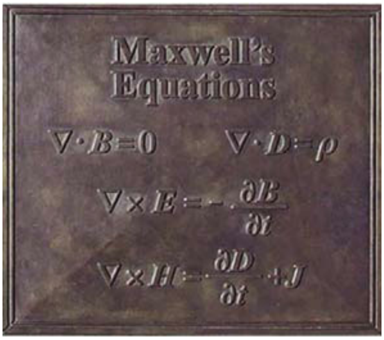


Fig. 2.10 Maxwell's equations

Chapter 3
Continuum Beyond Mechanics

3.1 Striking a Balance

Isn't it strange that humans, used to counting things and striking balances for millennia, came to the law of conservation of mass so late? The law of conservation of money is ancient (and became doubtful only when money became electronic), and even the alchemists aspired to transmute base metals into gold rather than to procure it *ex nihilo* as bitcoins are now mined. Of course, there could be ideological objections to conservation laws: God created the World out of nothing and has never since been constrained in working miracles. Those Greek and Indian philosophers who believed the world to be eternal stated that substance is permanent but, as usual, without any proof. The creation argument could be easily refuted by stating that only the Almighty is allowed to violate the laws He has set, but science provided no rational basis for conservation laws till the 18th century. Moreover, there was solid experimental evidence *against* conservation of mass: burned matter disappeared into thin air leaving just a heap of ash. Of course, Air was an Aristotelian element, and no one had any idea about its composition. On the other hand, the weight of metals increased as they were oxidized – another proof, though from the opposite side.

The earliest to proclaim the conservation of mass was Mikhail Lomonosov, the polymath once better known in Russia as a poet. In a letter of 1748 to Leonhard Euler, a fellow member of the Saint Petersburg Academy of Sciences[1], he stated: "All changes in nature are such that inasmuch is taken from one object insomuch is added to another. So, if the amount of matter decreases in one place, it increases elsewhere. This universal law of nature embraces laws of motion as well, for an object moving others by its own force in fact imparts to another object the force it loses". The law of conservation of mass was rediscovered and loudly proclaimed by Antoine Lavoisier (1789) who proved it (as Lomonosov reportedly did as well) by carefully weighing material oxidized in a closed vessel (Fig. 3.1). Lavoisier was guillotined in the reign of terror, among many others, being charged with financial machinations rather than

[1] It was later rephrased in his dissertation (Lomonosov, 1760).

L. Pismen, *The Swings of Science*,
https://doi.org/10.1007/978-3-319-99777-3_3

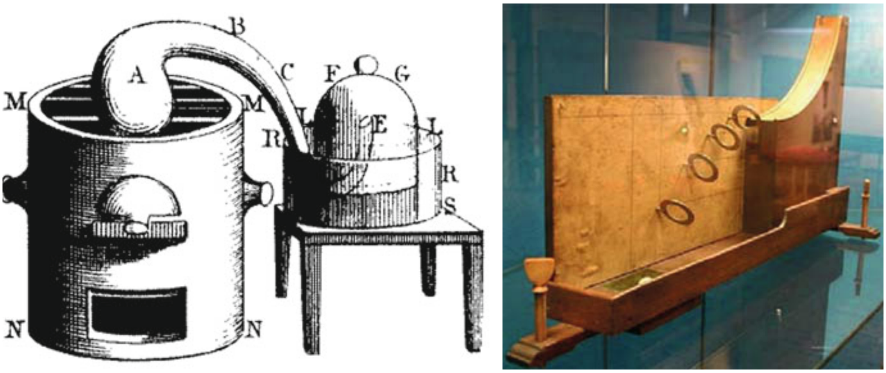


Fig. 3.1 *Left*: Lavoisier's oxidation experiment in a closed vessel. Engraving by Mme Lavoisier (Lavoisier, 1789). *Right*: Model of Gravesande's ball and clay experiment at the Boerhaave Museum, Leiden

with establishing this inconvenient anti-free-lunch law. The Republic had no need for scientists or chemists, as his judge observed.

The energy conservation law is more subtle, as energy exists in many forms, of which at least two, the kinetic energy of motion and heat, collided in the steam engines of the Industrial Revolution. These two forms of energy were, on the face of it, of totally different nature. Mechanical motion had obeyed Newton's laws since the 17th century, but they were centered on conservation of momentum rather than energy. Both, incidentally, were mixed up in the above quotation from Lomonosov. The importance of kinetic energy, called "*vis viva*" (proportional to the square of the velocity, while the momentum is linearly proportional to it), was suggested by Gottfried Wilhelm Leibniz and demonstrated more substantially by Willem Gravesande who dropped heavy balls into clay (Fig. 3.1 right) in 1722 (Hutton et al, 1809) and later by Émilie Du Châtelet, the first female physicist and the translator of Newton's *Principia*. However, kinetic energy is not conserved even in the absence of friction, and this might have misled Gravesande to support a fraudulent claim regarding a perpetual motion machine. Only the sum of kinetic and potential energy is conserved when there is no dissipation, as can be seen by following the swings of a pendulum; this is also the basis of Bernoulli's principle, which relates the change in hydrodynamic pressure to the flow velocity (see Sect. 2.2). The focus shifted from Newton's conservation of momentum to energy conservation in the Lagrangian and Hamiltonian formulations of classical mechanics.

There was no lack of ideas for a mechanical *perpetuum mobile*, some of them ingenious, some fraudulent, but none working (Fig. 3.2). Serious people would not believe this to be possible, especially as there was never any proof. Leonardo left a note in 1494: "O speculators about perpetual motion, how many vain chimeras have you created in the like quest? Go and take your place with the seekers after gold" (McCurdy, 1906). Some ideas that were not purely mechanical were based on the clandestine use of some resource, e.g., water evaporation. There were also mix-ups

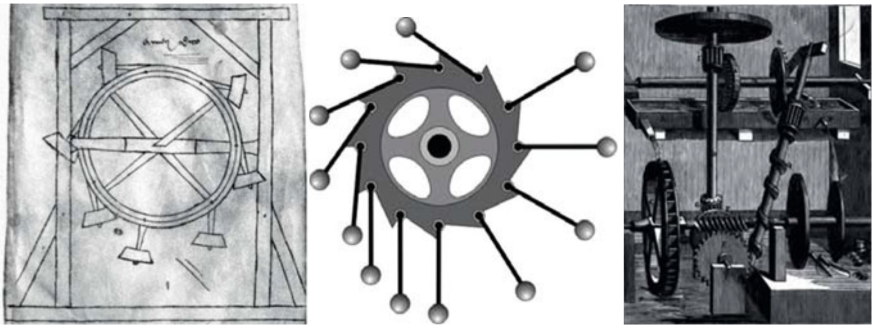


Fig. 3.2 Perpetual motion machines. *Left*: Perpetuum mobile of Villard de Honnecourt (about 1230). *Middle*: The overbalanced wheel. *Right*: Robert Fludd's water screw (1618)

between perpetual *motion* and perpetual *movers*. After all, the molecules of gas in an isolated container keep moving, it's just that their energy cannot be used without connecting them to the outside world. Electrons in atoms move as well, but their energy cannot be used if they reside in the lowest energy state. The recent idea of "time crystals", structures whose lowest energy states are periodic both in time and space (Wilczek, 2012a), is just another reiteration of motion that cannot be used for moving anything. On the ultimately large scale, the Universe at large is an example of perpetual motion which doesn't move anything.

In the 18th century, heat was still viewed as something unrelated to mechanics. Lavoisier came up with the idea of *caloric fluid* as the substance of heat. This fluid was supposed to be conserved and flow from hotter bodies to colder bodies. This was in line with a general attitude that can be traced back to Aristotle, assuming that any property of a material is determined by what this substance contains. The caloric theory was soon compromised by Benjamin Thomson (Rumford, 1876). First, he proved that materials do not change weight after heating or cooling, so that caloric fluid should be weightless. This was still not too bad, since such an ethereal substance could still exist. But then in 1798, something worse happened for the theory when he reported from his investigation of the heat produced while manufacturing cannons (Fig. 3.3) that friction could be an inexhaustible source of heat. This contradicted the "realistic" Lavoisier assumption that caloric fluid is a conserved substance, and it was a clear, though not yet quantitative, demonstration of the conversion of mechanical energy into thermal energy. But why was such a demonstration needed at the time? Steam engines had been already converting heat to mechanical energy for about a century (not counting Hero of Alexandria), and this could have given a hint that the two forms of energy were indeed mutually convertible.

While analyzing the efficiency of steam engines (still very low at the time), Sadi Carnot (1824) arrived at the idea of the Carnot cycle, the theoretical basis of the most efficient engine converting heat into work, which can also use mechanical work for refrigeration when operating in reverse. The operation of the Carnot cycle,

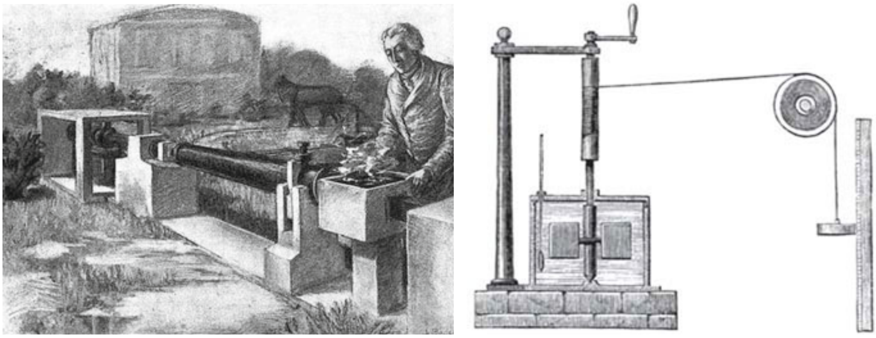


Fig. 3.3 *Left*: Rumford making water boil with heat generated by cannon boring (Hmolpedia). *Right*: Joule's apparatus for measuring the mechanical equivalent of heat

as it became clear later in his century, is based not only on the first law of thermodynamics (conservation of energy), but also on the second (an increase in entropy – see Sect. 4.1). Carnot, however, did not establish these laws – perhaps he would have if he had lived longer – and still relied on caloric theory. Following Faraday's discovery of magnetic induction and development of the electric motor, another form of energy entered the scene, and the mechanical equivalent of heat was found experimentally by James Joule (1845). By the mid-19th century, conservation of energy was firmly in place.

Einstein's $E = mc^2$ summed up the conservation of mass and energy, and even postmodern cosmologists note with satisfaction that our spacetime is flat, so that the total energy of the Universe, including also dark matter and energy, vanishes, at least approximately, and hope it to be exactly zero, so that even the act of creation, contrary to the Biblical account, was not in fact a free lunch. It would be difficult to prove conservation of mass–energy experimentally under absolutely all circumstances. There should be a general underlying principle sustaining it as a true universal law. This principle was formulated by Emmy Noether (1918), considered to be the greatest woman mathematician ever. Noether's first theorem states that any continuous symmetry of a physical system brings about a corresponding conservation law. Mass–energy conservation follows from invariance in time, while conservation of momentum follows from translational invariance. In relativity theory, these two conservation laws are brought together in the conservation of the four-dimensional energy–momentum tensor.

3.2 Transport Processes

With the conservation laws established, differential equations of heat and mass transfer could be put together. Joseph Fourier (1822) put forward the law of heat

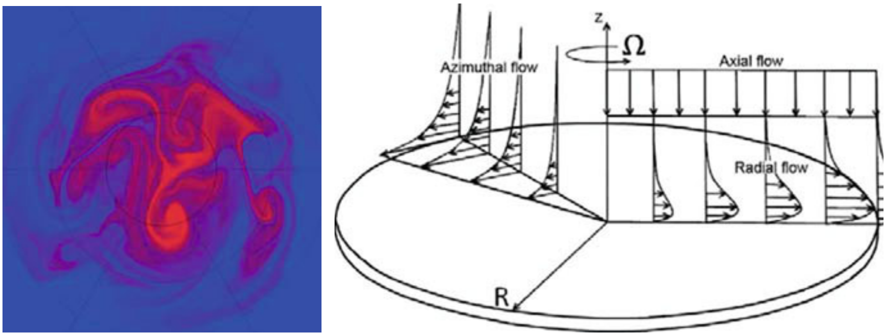


Fig. 3.4 *Left*: Chaotic mixing pattern. *Right*: Flow onto a rotating disk

conductivity modeled on Newton's linear relationship between the viscous stress and the velocity gradient: the heat flux is proportional to the gradient of temperature. The Fourier series, an indispensable analytical tool, was developed for the specific aim of solving the heat conduction equation. In a similar way, the physiologist Adolf Fick (1855) established Fick's law of diffusion. In the absence of chemical reactions, chemical species diffuse independently, each with its own diffusivity, which is the proportionality constant in the linear dependence of its flux on its concentration gradient.

Theories of heat and mass transport did not have such a glorious history as fluid mechanics. They were born later and quickly descended into engineering applications. The linear equations were too bland, but their extensions, combining heat and mass diffusion with flow, were too complicated for the taste of theoretical physicists and mathematicians. Heat and mass transport, like all dissipative processes, lack the rich mathematical structure of ideal fluid dynamics. Diffusion in liquids is slow: the diffusivity of a dissolved species in water is lower than the dynamic viscosity by at least three orders of magnitude. This is why we can restore a tainted blob by reversing the Stokes flow (see Sect. 2.3): the coloring ingredient is tardy to spread out, and the boundary of the colored domain is only slightly blurred after the reversal is completed.

This wide separation of scales has two consequences. The first is practical: it is hard to mix, for example, different chemical reactants, and all kinds of industrial mixers are used for this purpose. Complex flow geometries are engineered to enhance mixing of ingredients in polymer processing (Tadmor and Gogos, 2006). Salinity gradients in oceans persist, as do concentration gradients of nutrients and dissolved gases like oxygen and carbon dioxide. Powerful oceanic currents do not suffice to mix the waters, and even biogenic mixing by the drift of flocks of small marine animals may play a substantial role (Leshansky and Pismen, 2010). Chaotic advection patterns (Fig. 3.4) provide the best way to mix fluids efficiently (Aref, 2002).

The second consequence, most relevant for theory, was the necessity to study thin boundary layers. Walther Nernst (1904), a founding father of modern physical chemistry, considered diffusion through a stagnant fluid layer to be the slowest limiting step of chemical reactions on a solid surface. Transport can be accelerated by arranging for a flow pattern that would bring the dissolved reactants to the surface. An efficient device of this kind is a rotating disk sucking in the fluid and expelling it centrifugally (Fig. 3.4). The great advantage of this device is that the rate of mass transfer is uniform all over the disk, and this enables precise measurements of the rates of fast reactions, especially in electrochemical cells (Levich, 1962). Convective heat and mass transport are further intensified in turbulent flow. Both viscous and thermal or diffusional boundary layers of different thicknesses commonly arise, and the theory of turbulent mixing is even more complicated than the already intricate theory of turbulence (see Sect. 9.1).

Linear dependences between a flux and a driving force, like the laws of Newton, Fourier, and Fick, exist in other transport processes. One example is Ohm's law, relating electric current to voltage, i.e., electric potential difference, or, in the continuous formulation, the gradient of the potential. There are also cross-interactions between fluxes and driving forces of different nature. These include the Soret effect which is mass transport caused by a thermal gradient, the Dufour effect which is heat flux due to a concentration gradient, the Peltier effect which is heat flow in an electric field, the Seebeck effect which is an electric current caused by a temperature gradient. The diffusion of a certain substance can also be influenced by gradients of other chemical species. In a more precise formulation, the diffusional flux is proportional to the gradient, not just of the concentration, but of the chemical potential, which depends on the concentrations of all other species present, whence Fick's law is strictly valid only in dilute solutions.

Lars Onsager (1931) established reciprocal relations between different pairs of forces and fluxes, which imply that the matrix of phenomenological coefficients can always be transformed to a symmetric form. This ensures that the evolution toward equilibrium is monotonic – the basic law of non-equilibrium thermodynamics. At the time, his ground-breaking paper won him neither tenure at Brown University nor even a PhD, but only fame later on, with the 1968 Nobel Prize in Chemistry. Both linearity and the Onsager relations can be violated in systems maintained far from equilibrium by external forces or fluxes.

3.3 Chemical Reactions

In the Age of Enlightenment, fire, an Aristotelian element, was turned into a transformation process.

> And new Philosophy calls all in doubt,
> The element of fire is quite put out. (John Donne)

An Aristotelian atavism was still attributing flammability to a fluid contained in the flammable material. This fluid, called *phlogiston*, was supposed to be released during combustion. The phlogiston theory was so well established that Joseph Priestley, after discovering that oxygen sustained fire better than plain air, called it "dephlogisticated air". Phlogiston was quite a strange substance: since the weight of metals increased upon oxidation, it seemed to have negative mass. Carl Wilhelm Scheele, who discovered many elements, including also oxygen, which he called "fire air", did not dare to purge phlogiston. Henry Cavendish (1766), discovering hydrogen, called it "inflammable air" and thought that it might be phlogiston itself.

Lavoisier set everything the right way up, abolishing phlogiston (Lavoisier, 1783), calling oxygen oxygen and hydrogen hydrogen, establishing chemical nomenclature and stoichiometry, and launching rational chemistry (Lavoisier, 1789) which, 200 years on, would bring us fertilizers, poisonous gases, synthetic fibers, designer drugs, and more[2].

Fig. 3.5 Fire

Fire was now downgraded to just a chemical reaction, one of many. Like any reaction, it requires two conditions: supplying reactants (in this case, fuel and air) and maintaining a sufficiently high temperature to make the reaction rate sufficiently fast (see Sect. 4.5). Both tasks couple chemical reactions to transport processes. When the reactants are not mixed, the reaction takes place in a thin layer on the boundary between them; thus, wood burns from the surface, and so does oil on the tip of a wick. The supply of air is helped by flow – but it plays contradictory roles. In stagnant air, the supply of oxygen may not be sufficient, but if the advection is too fast, it may douse the flame by cooling it off. Thus, you fan the bonfire to start it, but blow a candle out. All this sounds simple and bland, but it is hard to express rationally. There is a busy line of research in "mathematical combustion" (Buckmaster and Ludford, 1982). It has been called mathematical, first, because it was initiated by applied mathematicians, and second, because it employed approximations that enable mathematical analysis but are not always applicable to real fires.

Heat and mass transport are particularly important in heterogeneous catalytic reactions. They require the reactant to be in contact with as large a catalytic surface as possible, and the most practical way to arrange it is to make a catalytic particle porous (see Sect. 3.5). However, this considerably slows down the supply of reactants and the removal of heat generated by exothermic reactions. In the 1930s, this problem and similar strongly nonlinear problems combining transport and chemical

[2] Fritz Haber, instrumental in the first two achievements in this list, won the 1918 Nobel Prize in Chemistry for the essential process of fixation of atmospheric nitrogen (right when his country lost the war). A Jewish German patriot, he was "thanked" by the Nazis in the last year of his life, but the Fritz Haber Institute that he founded remains one of the leading institutions of the Max Planck (formerly Kaiser Wilhelm) Society.

reactions attracted the attention of the young physicists Yakov Zeldovich (1985) and David Frank-Kamenetsky (1969). The references cited bear much later dates, but their early work was instrumental in the development of the Soviet nuclear bomb in the 1940s, more than in chemical engineering. Zeldovich, a self-taught genius, later became prominent in both nuclear physics and relativistic cosmology.

In the 1960s, the volatile Soviet leader Nikita Khrushchev embraced chemistry alongside planting corn and building drab housing blocks, now being torn down but offering quite an improvement over the communal flats we were living in prior to that. At the same time, cybernetics, denounced under Stalin as an imperialist ploy, was enthusiastically promoted. These developments worked together with the result was that computer labs were opened in chemical engineering research institutes. There were few people even remotely qualified to lead these labs, and I, as a half-baked PhD (Russian "candidate") was offered the job in a classified institution loosely connected with producing chemical weapons. I refused, but probably would have been rejected anyway after the relevant authorities had looked more closely at my file: besides having the wrong ethnicity, I mixed with the wrong kind of people. The job was taken by another young PhD who would become my friend; his was a similar background, but his questionnaire entries were more suitable.

At the time, we were quite enthusiastic, naively believing that we could solve the appropriate equations, with the rates of chemical reactions extracted from experimental data in a manner presaging the modern "Big Data" approach, the aim being to design chemical reactors *ab initio*, thereby escaping the usual engineering routine. All this, with the computers of the time, where the programmer had to write down explicitly which entry should be moved into which memory cell, perforate the cards, and wait, wait, wait until an error report came. While I was working on my PhD, an engineer colleague bet me that a particular result was wrong; I lost a dinner. I penned a sonnet some time before that about a gnome living in a computer and reprogramming it to encode his little tinkling bells. In modern computers, only bugs have space to move around.

Instead of leading a lab, I became, for half the salary, what would now be called a postdoc under Benjamin Levich, thereby acquiring the honor of being one of the multitude of Lev Landau's scientific grandchildren. Levich had his own agenda. Lifted by the success of his ground-breaking book (Levich, 1962) and praised by Western colleagues, he assumed that work on chemical engineering problems dealt with before that by mathematically illiterate engineers would secure him election as a full member of the Academy of Sciences. He summoned the theoretical department he chaired, some twenty physicists, PhDs and graduate students, working mostly on different kinds of electrochemical problems, from quantum theory to transport to neural activity, and commanded everybody to switch to chemical engineering for two years. For a time, I gave seminar talks almost every day on the related subjects, probably learning more in the process than educating others.

Later Levich "complained" that he had made a mistake: he meant to hire a chemical engineer, but hired a physicist. This was a compliment, but he did make a mistake: the authoritative figures in the field who also happened to have strong party connections did not tolerate the invasion. Levich was not elected and shortly afterwards applied for an exit visa, was refused, and, being deprived of his department, wilted away. Out of his natural milieu, he disappointed his Western admirers when he was eventually rescued by Edward Kennedy, but he kept prestigious appointments in New York and Tel Aviv. I was also hit and, having reached the glass ceiling, enjoyed writing popular science stuff for a while before moving away. In retrospect, living under the decaying Brezhnev regime in Moscow (which I craved to escape) was more fun than getting into an Israeli professor's rut, especially when young. I told the story about Zeldovich, Levich, and their friends and enemies in a paper entitled "The Golden Age of Chemical Engineering in Russia", prepared for an American Chemical Society meeting. It was not published, as, with *détente* in full bloom, the editors of the Proceedings judged it too inflammatory. I had more success in evading censorship in the USSR but, of course, I was more cautious there.

Prior to the time when Levich's assault on chemical engineering unfolded, similar developments took place across the ocean and on the other side of the Iron Curtain. Neal Amundson and Rutheford (Gus) Aris, both PhDs in mathematics with chemical engineering experience, successfully advanced mathematical modeling of chemical reactors in Minneapolis. Together with L.E. (Skip) Scriven, an outstanding physical chemist, and others, they turned the Chemical Engineering Department at the University of Minnesota into the world's best. One day, I was surprised by a letter from Dan Luss, a former student of Amundson and soon to be the Chairman of the Chemical Engineering Department in Houston who had dug out my paper on symmetry breaking (presented as his own by a Czech seminar speaker). Answering him was a mess, as an innocuous letter on its way abroad had to pass several levels of censorship. When, with the Iron Curtain rusting, Aris and Luss came to Moscow, they were told that I was vacationing while I was trying to contact them. Our world is small[3]: when I came to Technion, it turned out that Luss and the Department Chairman who hired me were brothers-in-law. We became good friends but our interests soon diverged. I told Aris while he drove me around Los Angeles that I had translated his book on chemical reactors into Russian – copyright observance in the USSR was in such a state that this was the first time he had heard of it.

From then on, everything calmed down. Computers became cleverer, and chemical engineers embraced them, often using wrong models, but correcting the actual design by tinkering. The dream of *ab initio* modeling turned out to be as evasive as the dream of "one principle sufficing to obtain everything from nothing" (recall Sect. 1.4). Physicists and mathematicians moved away to more interesting and less complicated problems. The pinnacle of my career in chemical engineering was a plenary lecture at the International Congress on Catalysis at the famous Negresco Hotel in Nice in 1980. As I spoke on bifurcations, I saw people leaving the lecture hall. The pinnacle stood at the edge of a precipice. Soon afterwards, my publications switched from Chemical Engineering Science to Physical Review.

[3] The term "small world network" is a mathematical one.

3.4 Charges and Currents

Charge transfer – electric current – is a no less important transport process, and it is also coupled with chemical reactions in electrochemical cells. In the beginning, electricity was static. Joseph Priestley (1767), experimenting with electrically charged spheres, conjectured that the force between charges falls off as the inverse square of the distance, similar to Newton's gravitation law. Others before him suspected the same relationship, which seems, by analogy, very natural. The law now bears the name of Charles-Augustin de Coulomb (1785), who measured the force precisely after inventing his torsion balance (see Sect. 2.4). The similarity is not coincidental. All *static* classical field problems are basically the same. You have a source: a gravitational mass, a charge, a magnet, a heat source, a tap and a sink – and a medium where it spreads – of any suitable nature, from vacuum to bath tub. Instead of a source, you may have a sink, which is just a source with a negative sign. The field generated by all kinds of sources and sinks is most easily characterized by a potential which satisfies, away from sources, one and the same Laplace equation, becoming the Poisson equation when sources are distributed. All measurable effects, such as the force exerted by the electric or magnetic field, the velocity of an ideal fluid, the heat or mass flux, and so on, are defined by the gradient of the appropriate potential. The inverse square law dependence is just a consequence of the dimensionality of space: in a two-dimensional flatland, the force would just be proportional to the inverse distance.

There are no negative masses – if there were, general relativity would bring about more bizarre spacetime structures, but there are positive and negative electric charges, north and south magnetic poles, and of course, it is possible to cool material as well as heat it. The decision about what to call positive or negative, as well as what to call north or south, is arbitrary. By chance, it was decided to say that the charge on glass rubbed by a cloth was positive, and hence by necessity, all charges repelled by this glass. The north magnetic pole, naturally, is the end of the magnetic needle pointing to the north, something the Chinese were using for navigation since the 11th century and which Europeans only caught up with more than a century later. Both designations turned out to be inverted. Electric current is carried through metallic conductors by negatively charged electrons, about which nobody could have known before electrons were discovered by J. J. Thomson (1897), late enough to earn him the 1906 Nobel Prize in Physics. As a result, the conventional direction of currents in electric circuits is opposite to the actual electron flow. Moreover, since opposite magnetic poles attract, the magnetic pole in the Earth's Northern Hemisphere is actually a south magnetic pole.

Static electricity and magnetism have been known since the depths of antiquity. Presaging Maxwell, Thales of Miletus unified electricity and magnetism, erroneously assuming them to be of the same nature, as he unified all elements in Water. The handling of static electricity was greatly advanced in the 17th century by generating strong charges with the help of friction machines and storing them in a Leiden jar, the original form of a capacitor. On the other hand, magnetism had always had something magical about it, as it was independent of our actions, such

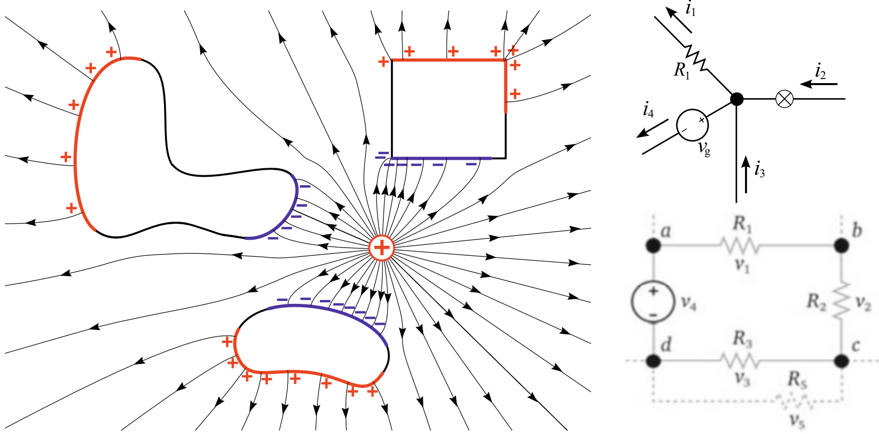


Fig. 3.6 *Left*: Electrostatic induction of charges on conductors by the electrostatic field (depicted by *lines with arrows*) of a positive charge. *Right*: Kirchhoff's circuit laws. *Above*: The current entering any junction is equal to the current leaving that junction. *Below*: The sum of all the voltages around a loop is equal to zero

as rubbing something to induce electric charge. Thales viewed the magnet as alive because of its attractive powers. The magic survives in its figurative meaning as the ability to attract and charm people. Even now that we know how to induce magnetic fields, people still buy magnetic bracelets with alleged healing powers. Perhaps, the impossibility of separating opposite magnetic poles, unlike positive and negative electric charges, contributed to the enchantment. The magnetic monopole, a hypothetical elementary particle predicted by Paul Dirac (1931), was never found.

Static electric and magnetic fields, as well as steady currents, were well understood by the early 19th century. A steady current in a conducting wire was defined by Ohm's law to be proportional to the applied voltage (the difference between the potentials at its ends); and this simple law was generalized to a network of conductors by Gustav Kirchhoff in 1845, while he was still a student – the same Kirchhoff who later contributed to elasticity theory, spectroscopy, and other fields. Henry Cavendish reportedly discovered both Coulomb's and Ohm's laws, but being a true natural philosopher, rich and reclusive, did not care to publish. Besides the forces, an important result of the analytical theory was understanding the distribution of charges on the surface of conducting bodies (see Fig 3.6). More mathematical elegance was added by Gauss's law (presaged by Lagrange some forty years earlier), which relates charges in a certain volume to fluxes through the surrounding surface. Laws are not always called by the names of their earliest discoverers, as those coming later may better formulate and/or better advertise them.

Things became more complicated again as non-stationary processes appeared on the scene. Non-stationary equations have more variety than static ones. In the simplest case, the Laplace equation containing the sum of second spatial derivatives can be supplemented by either the first or the second time derivative. The results

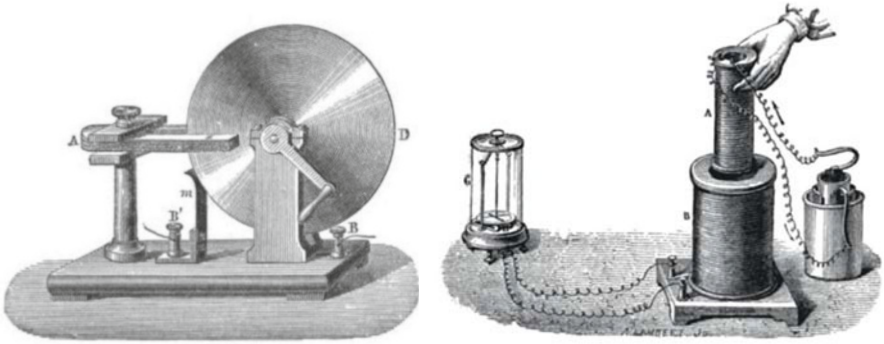


Fig. 3.7 *Left*: A Faraday disk (D) rotated in the magnetic field of a horseshoe magnet (A) induces an electric current in the radial direction; the electric circuit is closed by a wire connecting the sliding contact (m) with the axis. *Right*: The magnetic field of the electric current through the moving coil (A) induces a current in the large coil (B)

could not be more different. In the first case, we have an irreversible dissipative process, as in heat conduction or diffusion. In the second, what comes out is the non-dissipative wave equation.

Non-stationarity unified electric and magnetic phenomena. It started on an empirical level in 1820, when Hans Christian Ørsted discovered by chance that switching the electric current from a battery on and off caused a compass needle to orient itself perpendicularly to the conducting wire. In the same year, a week after hearing of this discovery, André Marie Ampère demonstrated that parallel currents repel each other and opposite currents attract as magnetic poles do, and Jean-Baptiste Biot and Felix Savart showed that the magnetic force exerted on a magnetic pole by a wire is oriented perpendicularly to the wire and falls off inversely proportionally to the distance, in accordance with Coulomb's and Gauss's laws. Soon afterwards, Michael Faraday burst onto the scene. Buoyed by intuition rather than mathematical theory, he invented the electric motor in 1821 and in 1831 discovered the electromagnetic induction of a current in one circuit by the current in a neighboring moving circuit (Faraday, 1839), which led to the construction of the transformer and the electric dynamo (Fig. 3.7). The 19th century advance culminated, on the one hand, in the revolution by Maxwell (1873) that unified electromagnetism with optics (see Sect. 2.5) and, on the other hand, in the rapid development of electric power generation and usage, initiated in the late 19th century by Thomas Alva Edison and Nikola Tesla. Faraday could have scooped Edison, but applying scientific knowledge to personal enrichment was below his dignity.

Another unification movement developing at the same time connected electricity with chemistry. It started with the first electric battery invented in 1793 by Alessandro Volta. A basic battery uses two electrodes made of different metals placed in an electrolyte (an ionic conducting liquid, such as salt water). The metal with a lower electrochemical potential (such as zinc relative to copper in Volta's cell) dissolves,

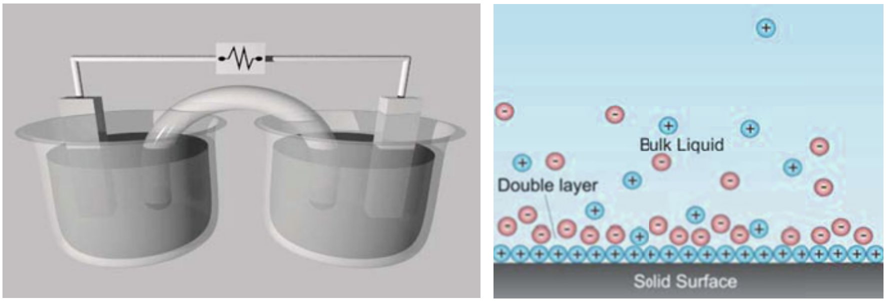


Fig. 3.8 *Left*: A basic electrochemical cell. *Right*: The Debye layer

and the departing positively charged ions leave electrons on the electrode; they are subsequently conducted to another electrode which acts just as a supporting actor in this show (see Fig. 3.8). Of course, Volta could have had no idea of all this. It took Humphrey Davy (who later employed a bookbinder's apprentice Faraday as his assistant) to explain that Volta's cell is based on chemical action. The reverse process of transmuting electric to chemical energy was first carried out in 1785 by Martin van Marum using voltage supplied by an electrostatic generator. This process, later called *electrolysis* by Faraday, was instrumental in the discovery of highly reactive elements, such as alkaline and alkaline earth metals isolated by Humphrey Davy, and it became industrially important in the 20th century when mechanically generated electric power became abundantly available.

Davy himself did not yet know either about electrons or about ions. Further progress once again required Faraday who, after finishing with electromagnetism, turned in 1833 to electrochemistry. Faraday called carriers of electric charges *ions*, a derivative of the Greek verb "to go": there were positively charged cations and negatively charged anions. Atoms and molecules were still hypothetical entities at the time, though realistically minded scientists quietly adopted them as intuition aids. In 1884, Svante Arrhenius submitted a dissertation explaining that the dissociation of dissolved salts into positive and negative ions was the basis for electrolytic conductivity. It brought him the 1903 Nobel Prize in Chemistry, one of the first awarded, though it was not his most important contribution to physical chemistry (see Sect. 4.5).

Studies of ionic transport and electrochemical reactions became a lively field of research in the 20th century. The behavior of ions in electrolytes became better understood through the work of Peter Debye; in particular, he elucidated the structure of double layers near electrodes with a dense sheath of ions of the opposite sign and a diffuse layer of ions of the same sign (Fig. 3.8). The diffuse layer spreads out as the concentration of ions decreases. Paradoxically, this becomes problematic in dielectrics, which are usually contaminated by small amounts of ions. As such a fluid flows through a tube, ions are flushed out and charges accumulate at the exit; this led to sudden explosions before the cause was understood.

The kinetics of electrode reactions is the most difficult electrochemical problem, as it may be still more complicated than the kinetics of other surface processes, such as adsorption and catalysis (see Sect. 4.5). A prominent role in the study of kinetics was played by Alexander Frumkin at the Institute of Physical Chemistry in Moscow that now bears his name. One of the most important modern industrial projects is the construction of powerful and light-weight rechargeable batteries and accumulators for all kinds of devices from phones to cars, and also fuel cells that could power cars without polluting the atmosphere. Modern electrochemical cells are very much unlike those of Volta, Faraday, and Debye, and commonly use solid electrolytes.

3.5 Porous and Granular Media

Continuum equations, the best tool for an elegant mathematical analysis, were extended to materials not looking continuous at all – porous and granular solids. In the spirit of all simple linear laws, Henry Darcy (1856) deduced the law bearing his name from experiments: the flow velocity is proportional to the pressure difference (or gradient, in the differential formulation). This law, rather similar to Ohm's law with the current replaced by the velocity and the voltage by the pressure, is simpler than the hydrodynamic equations, and even simpler than the Stokes equation of viscous flow. This simplicity masks the complexity of the internal structure of porous media, hidden in the empirical coefficient of this linear dependence. Like any phenomenological dependence, it is valid at distances that are large compared to a microscopic scale where the medium cannot be treated as a homogeneous continuum. For Newton's viscosity or Hooke's elasticity laws, this scale is molecular, but for Darcy's law, it is macroscopic, corresponding to the thickness of the pores and interstices. Moreover, pores are irregular, tortuous, and corrugated. Computing permeability – the coefficient of Darcy's law – requires far-reaching assumptions about the internal structure, and even then it is a tortuous task. This is a common story. Computing the viscosity of liquids on the basis of their molecular properties is just as hard and unreliable.

Besides the permeability, we need to know the diffusivity, which is reduced compared to the value in the bulk fluid. We need to know it still more precisely, since, in contrast to what happens in the bulk, it is not helped by advection, this being still more strongly suppressed. This is particularly important for heterogeneous catalytic reactions (see Sect. 3.3) which are carried out inside porous particles to pack a lot of surface into a small volume[4]. We once suggested (Nir and Pismen, 1977) organizing a catalytic particle like an urban road network, with highways wide enough to enable advection to penetrate the network of narrow capillaries with an abundant surface area. The hierarchy goes further down: from relatively thick pores with diameters far exceeding the molecular mean free path to exceedingly narrow ones where diffusing molecules collide with walls more often than among themselves.

[4] A millimeter-sized grain may contain as much area as a middle-sized farm – this was used as a lively image in advertisements.

Porous materials hierarchically structured at small scales are common in nature, and modern technology aspires to imitate them (Su et al, 2012).

More complications arise when gas and liquid coexist within pores. Then capillary forces intervene (see Sect. 5.3) and if the liquid wets the solid[5], it tends to accumulate in narrower pores. A chemical reaction induces composition changes across the porous body; as a result, this affects the evaporation equilibrium, and the vapor evolving in high volatility regions moves through gas-filled pores, condensing in low volatility domains. Capillary forces drive the condensed liquid back, thereby generating countercurrent gas and liquid flows that are able to reduce concentration gradients in the liquid phase much more efficiently than slow diffusion. If the liquid is conducting, only its connected clusters contribute to the electric current, and the conductivity drops to zero if no connected liquid-filled network exists (Fig. 3.9). The formation of such a network is a critical event, attained at a certain filling ratio that depends on the network structure. This is the subject of percolation theory (Stauffer and Aharony, 1992), which has many other applications, from solid state magnetism to communication networks. Percolation thresholds are known for a number of regular lattices, but not for a random network of irregular pores. Transport in porous media partially filled by liquids is also important for other practical applications. This is what determines the distribution and motion of water and nutrients in soils and the success or failure of oil extraction from oil sands.

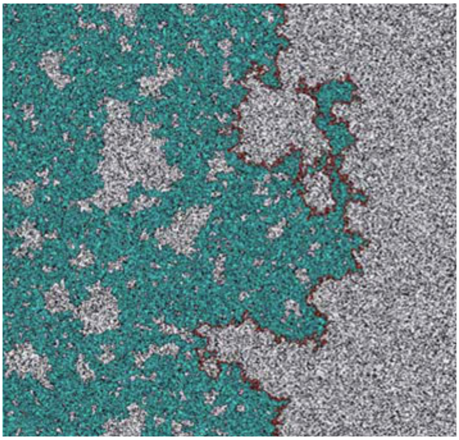


Fig. 3.9 Percolation front near a percolation threshold

The problem further complicates when the material of the porous body is flexible, as in a sponge. Such media were first studied in the geophysical context by Maurice Biot (1941), who pioneered the field of *poroelasticity*. The study of this kind of medium requires us to combine the hydrodynamics of liquid flow with the elasticity of a flexible solid material. In geology, this involves processes unfolding on long time scales under the action of strong forces and then unleashed in violent earthquakes. Far gentler processes that attract the most attention in our day take place in live cells and tissues: their deformation involves the motion of both the viscoelastic active cytoskeleton and the liquid cytoplasm (Moeendarbary et al, 2013). Intracellular dynamics is further complicated by the various chemical transformations in this crowded environment (see Sect. 9.2).

[5] Otherwise it would not penetrate – but the distribution of pore sizes is often determined by pushing a non-wetting liquid, usually mercury, into the pores.

Fig. 3.10 *Upper left*: Sand blowing off a crest in the Kelso Dunes of the Mojave Desert (photo by M. A. Wilson). *Lower left*: Sand wrinkles in shallow water (photo by the author). *Right*: A powder snow avalanche in the Himalayas (by Chagai)

Granular media are porous as well, but consist of unconsolidated solid particles, and can therefore flow. Microscopic solid particles move freely in dilute suspensions, and larger particles hover in fluidized beds sustained by upward flow[6] or go down by gravity in grain elevators. Sand grains are moved by wind or sea currents, forming a variety of patterns (Fig. 3.10). A granular pile may rest in equilibrium supported by force chains between the grains – but if the slope exceeds a critical incline at some location, an avalanche develops, which may be as huge as the one in Fig. 3.10 (right) or as small as in a child's playground. If snow falls or you pour sand on a sandpile or wind blows in the desert, a snow drift or a pile or a dune grows till the critical incline is reached, and the critical incline is restored following an avalanche. This is an example of self-organized criticality (Bak, 1996): the system tends to adhere to the brink of a disaster.

[6] This is an efficient way to enhance heat and mass transport in chemical reactions. The electrochemist Martin Fleischmann, later notorious for his cold fusion claim (see Sect. 10.3), used a fluidized bed of conducting particles to transfer charge between electrodes.

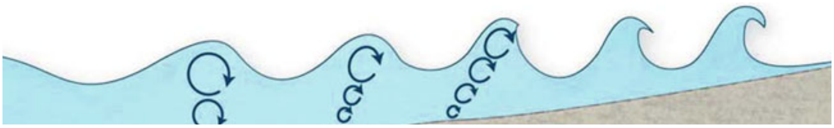


Fig. 3.11 Breaking waves

3.6 The Continuum Breaks Down

Solutions of continuum equations are not always smooth. I already mentioned boundary layers developing near solid surfaces in Sect. 2.5 and breaking waves in Sect. 2.3. In both cases, solutions of simple equations break down at some locations. In the former case, the ideal fluid equations are incompatible with the no-slip boundary condition enforcing zero velocity at a solid surface, and viscosity has to be taken into account to resolve this contradiction. In the latter case, the amplitude of the surface wave becomes comparable with the water depth, and the shallow water approximation can no longer be used; the waves steepen, and eventually break as the surface elevation ceases to be a single-valued function of the horizontal coordinate (Fig. 3.11). Boundary layers can also develop within the bulk of the fluid (Helmholtz's vortex sheet) or in the wake of a blunt body (von Kármán, 1921). Here again, sharp gradients are smoothed out by viscosity, but such boundary layers are apt to develop instabilities and eventually turn turbulent.

The failure of dimensional reduction that occurs in shallow water happens also in strongly deformed elastic sheets, such as crumpled paper (Witten, 2007). Mechanical stress tends to focus at some locations where sharp folds appear and the thin film approximation fails (Fig. 3.12). Singularities appear at vertices where these folds converge, and they can only be resolved by solving the full elastic equations at folds and vertices and matching them with smooth solutions elsewhere. L. Mahadevan FRS, the soft matter wizard at Harvard University, was awarded the 2007 Ig Nobel Prize for Physics for a theoretical study of how sheets become wrinkled. He proudly came to the ceremony to receive the award[7].

Fig. 3.12 Crumpled paper

In all these cases, singularities are resolved by adding more terms into the continuum equations. It may happen, however, that the singularity cannot be healed up to the atomic scale, or at least up to the scale of intermolecular forces. This is what happens at the mo-

[7] At some point, this prize became quite prestigious, with Joseph Keller, the Wolf Prize laureate (twice), Michael Berry FRS (also Wolf Prize), and Raymond Goldstein FRS among the winners. Andre Geim FRS is so far the only one to receive both the Ig Nobel (2000) and the real thing (2010)

ment when a liquid jet breaks into droplets, and even when a droplet slides across a solid surface (more on this in Sect. 5.4). Studying the asymptotic shape near the singularity can be revealing, but numerical computations are frustrating, as the result remains dependent on the mesh size until molecular scales are reached. This dilemma is still more frustrating in astrophysics. There is a “desert” of many orders of magnitude (recall Sect. 1.3) between “phenomenological” scales where all known physics unfolds and the Planck scale where a singularity, such as exists in the middle of a black hole or at the Big Bang itself, may or may not be resolved – if it is not resolved somewhere else on the way across this uncharted desert.

Back on Earth, some time in the late 1960s, the huge lecture theater of Moscow State University was overcrowded with lecturers and students who came to listen to the dispute between two prominent members of the Academy of Science: the applied mathematician Leonid Sedov and the physicist Yakov Zeldovich. The dispute had political undertones: mathematicians were largely ethnic Russians, communists, and in some cases antisemitic, while among physicists there were disproportionate numbers of Jews and latent dissidents. Sedov, together with the virulently antisemitic mathematician Pontryagin, was acting to suppress Zeldovich’s informal book “Mathematics for Beginners”. The subject was the shape of the tip of a propagating fracture, which Zeldovich thought to be rounded and Sedov, a sharp singularity. Physicists abhor singularities and always look for some factor working on a shorter scale that helps to close the problem. I do not remember who won at that time, or whether it just depended on your inclinations – but in the end both might be right: not all fractures are equal, the tip may be rounded by plastic deformation due to concentrated stress, but it may also be sharp up to the atomic scale.

Chapter 4
From Continuum to Atoms

4.1 Just a Hypothesis?

Is matter infinitely divisible or there are ultimate indivisible elements – atoms? Philosophers and scientists after them were of different opinions about this, and no definite proof appeared until the turn of the 20th century. The atomism of Democritus had a more materialistic flavor, but the problem with atoms was that they were supposed to move in the void, which many found objectionable. Plato disliked Democritus, but he might have meant in the obscure dialog "Timeus" that the elements consisted of the associated polyhedra (tetrahedron – fire, cube – earth, octahedron – air, icosahedron – water, and dodecahedron "arranging the constellations on the whole heaven"). He went still further suggesting that triangles, of which all Platonic bodies can be built, are the sole basic structure. Aristotle abhorred the void and thought matter to be infinitely divisible. Since he was the ultimate authority, and what was copied and therefore has survived was his and Plato's work, no wonder that everything known of ancient atomists comes from fragments and retellings.

Closer to our own age, in his dissertation "Elements of Mathematical Chemistry" written in 1741, Lomonosov gives the following definition: "An element is a part of a body that does not consist of any other smaller and different bodies [...] a corpuscle is a collection of elements forming one small mass." Dalton's chemical elements (Fig. 1.5) could be naturally associated with atoms. Laplace believed that natural phenomena could be explained in terms of interactions between particles and their evolution predicted by his demon (see Sect. 8.1). Stokes initially thought he could attribute viscosity to friction between some kind of particles. All this lacked proof. Heat could be viewed either as a fluid which could pass through continuous matter or a vibratory motion of molecules. The Royal Society twice rejected papers on the kinetic theory of gases, by John Herapath in 1820 and John James Waterston in 1845, as nonsense (Truesdell, 1980).

In his thick and dull history of science, John Bernal (1954) wrote that "the idea of atoms had seemed to be a revolutionary one and had always been associated with general revolutionary and atheistic thought". This was the official communist line,

L. Pismen, *The Swings of Science*,
https://doi.org/10.1007/978-3-319-99777-3_4

but it hardly goes further in history than Lenin's lambasting Ernst Mach while idling his time away in the British Museum between revolutions; this was in 1908, when the question was already decided. The distinctions were not so clear-cut, however. Newton was an atomist in his theory of light but his calculus of infinitesimals developed into an elegant theory of continuum mechanics; electromagnetism was too precious to be allowed to disintegrate into atoms; and Mach stood firm until the last. Newton tried unconvincingly to deduce Boyle's law of inverse proportionality of pressure to volume by assuming the interaction between particles to be inversely proportional to distance. The young Leonhard Euler (1729), though contributing to progress in the mathematics of continua, thought of air as consisting of whirling close-packed particles, and arrived at a coefficient for Boyle's law that was proportional to the square of their velocity – this was actually the opposite of Bernoulli's law of pressure decreasing with velocity (see Fig. 2.1).

The mid-19th century saw the development of thermodynamics. I could not understand while an undergraduate why the boring freshman course started with the Carnot cycle (Fig 4.1) that we have already met in Sect. 3.1. I couldn't care less about the efficiency of steam engines depending on the temperature difference between the hot and cold reservoirs. The Moscow river never froze in winter because of the effluence of water condensed at the low-temperature end of the Carnot cycle by power stations, and apartments were comfortably heated by the same effluence. The reverse Carnot cycle was still less relevant, as there were no air conditioners, and even refrigerators were rare.

The true reason why a traditionally designed course started with the Carnot cycle was the work of Rudolf Clausius (1850), who used it as a basic example for establishing the Second Law of Thermodynamics. Putting things simply, you cannot run the Carnot cycle back and forth without pumping in more energy. Clausius (1865) went on to introduce the notion of *entropy*, which, like our age, never decreases. His brief summary of the basic laws of Nature was: "The energy of the World is constant. The entropy of the World strives for a maximum." Entropy is not an intuitive concept and cannot be measured experimentally. Clausius only defined a *change* of entropy through the ratio of transferred heat to the absolute temperature – but the meaning of entropy could only really be understood in the framework of statistical physics, as we shall soon see.

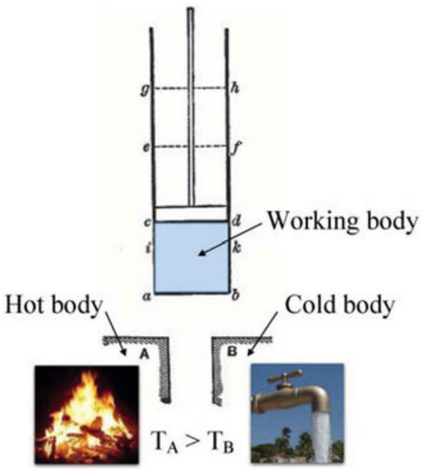


Fig. 4.1 The Carnot cycle

After finding what looked like a reasonable textbook, I decided that thermodynamics was all about partial derivatives. There are variables, like temperature, pressure, and volume (or density), and there are equations of state relating them for a particular substance. The earliest one, which said that the pressure of an ideal gas is inversely proportional to volume and went back to experiments by Robert Boyle (1662), was generalized to the law of ideal gases by Émile Clapeyron (1834), stating that the product of the volume and the pressure is proportional to the temperature. You can use these relations to differentiate one variable with respect to the other, keeping other variables constant, and get appropriate relations: compressibility or heat capacity or the coefficient of thermal expansion. Once you go to dense gases or liquids, these relations become more complicated, but they can be found experimentally as well. The situation becomes more difficult when there are phase transitions. In this case, the Clausius–Clapeyron relation can help us to draw phase diagrams like the one in Fig 5.5. You only need to know, again experimentally, specific heats and entropy changes at the phase transitions: evaporation of a fluid, freezing of a liquid, and sublimation of a solid.

You can go further and construct thermodynamic potentials: the Helmholtz free energy that has to decrease in a closed system evolving at a constant temperature and volume, and the Gibbs free energy that has to decrease in a closed system evolving at a constant temperature and pressure. Willard Gibbs (1876) brought thermodynamics to its highest point, extending it also to systems with variable chemical composition and introducing chemical potentials that govern the transport of different chemical species and the equilibrium of chemical reactions within a closed system, or mass exchange of an open system with its environment. Gibbs' masterpiece, full of long mathematical formulas and published in an obscure journal in the still provincial United States, was only noticed when Wilhelm Ostwald translated it into German and hailed Gibbs as the "founder of chemical energetics".

Einstein said that classical thermodynamics "is the only physical theory of universal content which I am convinced will never be overthrown". This universality comes at a price. There is nothing in this theory that clarifies the constitutive relations that are necessary for practical application of the thermodynamic laws, and experiment remains the only expedient. The two great laws, as well as the machinery of thermodynamic potentials, are so universal that they do not even depend on whether atoms and molecules actually exist, whence Ostwald, enemy of the atomic hypothesis, could hail Gibbs.

In the mid-20th century, there was a drive to extend the universality of classical thermodynamics (which should more properly be called thermostatics) to non-equilibrium processes (De Groot and Mazur, 1962; Prigogine, 1955). Non-equilibrium thermodynamics is largely based on Onsager's reciprocal relations between different thermodynamic forces and fluxes (Onsager, 1931). It is applicable, however, only close to equilibrium where linear transport laws apply. Chemical reaction rates already become nonlinear rather close to equilibrium, and in many cases (recall fire) equilibrium either does not exist or is trivial (e.g., completely burnt out). The most interesting phenomena are nonlinear and non-universal.

4.2 Molecular Kinetics

It was impossible to derive constitutive relations, even in simple cases, without considering the motions of atoms and molecules. What gives a tangible body to thermodynamics is *statistical mechanics*. Here comes James Clerk Maxwell (1867) again. He considered molecules of gas in constant motion, colliding with each other elastically and exchanging their kinetic energy. It is because of these collisions that diffusion is slow and the gas as a whole moves much more slowly than its individual molecules. Absolute disorder readily submits to mathematical analysis if you are interested in statistics rather than in the individual destinies of molecules or card players, and games of chance prompted the development of probability theory, starting from the 16th century. In ideal gases, where interactions between molecules are negligible, the velocity distribution depends only on the ratio of kinetic energy, proportional to the velocity squared, to thermal energy, proportional to the temperature. For each molecule, the probability of having a certain velocity decreases exponentially with this ratio. This is a particular case of the general law derived later by Gibbs (1902). Since the number of molecules with a certain velocity is also proportional to the velocity squared, the velocity distribution has a maximum, as shown in Fig. 4.2. As the temperature grows, the distribution flattens and its maximum shifts to higher velocities; the same happens when the mass of the molecule, and hence also its kinetic energy, decreases.

The velocity distribution is rather wide, and there are always "hot" molecules with velocities considerably higher than average. Maxwell discussed with William Thomson (Lord Kelvin) the possibility that a tiny creature, called by Thomson "Maxwell's demon", could open a shutter to let hot molecules from the left into the right-hand container in the inset of Fig. 4.2, while closing it to keep the others in, and letting less energetic molecules pass from the right to the left to keep the numbers even. Clearly, this would create a temperature difference between the two containers which could be exploited to produce work, in defiance of Clausius's

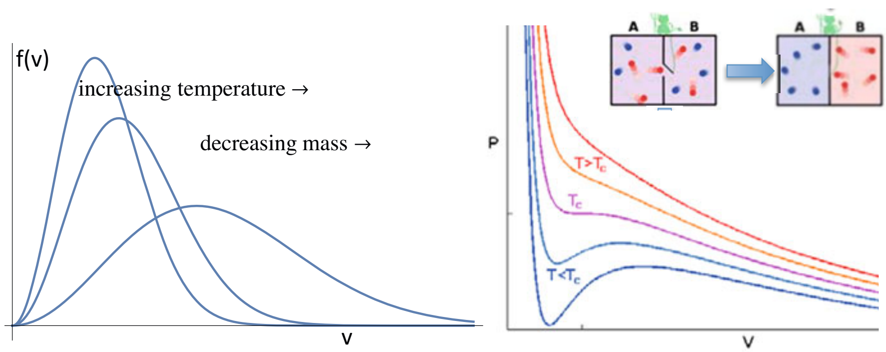


Fig. 4.2 *Left*: Maxwell distribution. *Right*: Van der Waals isotherms. Note the changing shapes of the curves below the critical temperature T_c. *Inset*: Maxwell's demon

motto. Neither Maxwell nor Thomson seriously believed in this proposal, but it initiated a prolonged discussion, which even became sort of practical in the 21st century when nanotechnology became capable of building real demons (Serreli et al, 2007). However, the real physical device required power to run it and therefore did not violate the Second Law. The word "demon" implies power and ill will, and it would have been more appropriate to call it instead an *imp*. A tiny imp should be a molecular-sized creature, which is also subject to fluctuations and would tend to attain equilibrium with its environment, that is, with the very same molecules it has to sort out. It should also be able to measure the velocities of passing molecules, and would need to create, store, and erase information. Information is the opposite of entropy, and handling it requires work.

Random motion of molecules easily explains the laws of ideal gases. The young Johannes van der Waals (1873) went further. He introduced a simple quadratic interaction between molecules and took account of their finite size. This allowed him to obtain a dependence of pressure on volume at a constant temperature (called an isotherm) which is not just a monotonic function but has a minimum and a maximum (Fig. 4.2). This means that the volume or density can be different at the same pressure: the lowest volume corresponds to the liquid, and the highest to the gas phase, while the middle point is unstable. The dependence becomes monotonic again above the critical temperature. This work was cited in the justification for his 1910 Nobel Prize in Physics, though he produced much more in the intervening years.

As Maxwell went on with the electromagnetic field, molecular kinetics was taken over by Ludwig Boltzmann (1897). His most famous statement, engraved on his tombstone[1] (Fig. 4.3), is that entropy is proportional to the logarithm of the number of ways of realizing a particular state. The constant *k* in this formula came to be called the Boltzmann constant. If there is only one possibility, the probability is one and the entropy is zero. This happens in a one-component perfect crystal at absolute zero temperature. Walther Nernst (1920 Nobel Prize in Chemistry) proclaimed it as the "Third Law of Thermodynamics" but it hardly deserves such a distinction. Boltzmann's formula is not universal: it is applicable only to an *ergodic* system that evolves over time to explore all accessible states. Boltzmann computed probabilities by discretizing the variables that characterize the state of molecules (positions and velocities) in "cells" of a finite size. This procedure becomes more natural in the framework of quantum theory, where quantum states are naturally discrete. The larger the number of cells with a certain value of

Fig. 4.3 Boltzmann's tombstone inscribed with his famous formula

[1] In a way, it had driven him into the grave.

a certain variable, the higher the probability that this value will be observed. There is only one cell with zero absolute velocity at a particular point, but at finite velocities the number is proportional in three dimensions to the velocity squared, and at large velocities the probability goes down, exactly like in the Maxwell distribution. Regarding the position, the probability of a molecule being within a certain cell is higher in a small closed vessel than in a large one. The latter example most easily demonstrates why probability is related to thermodynamic entropy. When kids are let out into a courtyard they disperse swiftly, but it takes discipline and persuasion to gather them back into the class. Molecules spread readily in the available space as well, but they are deaf to persuasion; in order to collect them in a small volume you need compression, and this requires work.

4.3 Entropy and the Arrow of Time

Maxwell's and Boltzmann's arguments were criticized as not being rigorous, and some counterexamples were even suggested, assuming some special initial arrangements of the molecules or violation of ergodicity. Controversies surrounding the molecular kinetics are discussed at length by Jos Uffink (2007). Most of them are now difficult to understand. The basic statements of molecular dynamics are justified by the purity of complete disorder in a huge molecular ensemble. More serious objections concerned the emergence of improbable states. Henri Poincaré (1890) proved the recurrence theorem stating that any non-dissipative mechanical system returns to its original state after a sufficiently long time. Based on this, he observed that "to see heat pass from a cold body into a warm one, it will not be necessary to have the acute vision, the intelligence and the dexterity of Maxwell's demon; it will suffice to have a little patience". Quite a lot of patience would be needed. Boltzmann came up with a calculation: if you take a billion molecules (which is still a billion times less than the number of molecules in a cubic centimeter of the air we breathe), the number of different ways in which the molecules can be distributed over cells about a nanometer size in the physical space and a meter per second in the velocity space is roughly ten to the billionth power, so you would have to wait about ten to the billion units of time – intervals between molecular collisions or seconds or days or years, your choice of time unit will not affect the result much. Numbers of this kind are not astronomical – they are out of this world.

This still didn't solve the problem, because at this time scientists assumed the Universe to be eternal. A still more outrageous idea was "Boltzmann's brain" arising spontaneously as a result of a giant fluctuation. I wonder whether it appeared to those involved in the discussion, including Boltzmann himself, that Boltzmann's brain did exist, along with many others. We could identify this giant fluctuation with our Big Bang, but conscious beings evolved in just a few billion years as a result of a rational evolution process, helped just a bit by random mutations. In some ways, the notion of a recurrence is optimistic. After the entropic death predicted by thermodynamics, one can wait "a little", and a fluctuation will set everything in motion again.

Unlike Maxwell who died early, Boltzmann was still alive when Nobel Prizes were being awarded; getting one would very likely have saved him from hanging himself – but his theories were still controversial, and he was frustrated by the unending struggle with the deniers of atoms, who never seemed to give up. One often needs to live long to get a Nobel Prize; it may honor work done 30 years before, and in the case of the 2003 and 2013 Prizes in Physics, almost half-a-century back. Still, though there have been a few controversies, Nobel Prizes in Physics and Chemistry are not as arbitrary as those in Literature and, still less than those distributed in the name of Peace. Anyone who follows Nobel Prizes will have examples, some too obvious to mention.

Claude Shannon (1948) inverted the sign of the Boltzmann equation to replace entropy by information. This version is apparently more popular than Boltzmann's original, since many people think they understand better what information is than what entropy is. In fact, information is more difficult to measure. Shannon, as a communications engineer, cared for the speed and capacity available for transmitting bits of information – but applying Boltzmann's formula to real information content leads nowhere. A book machine-produced in Jonathan Swift's Laputan Academy, or stored in Jorge Borges' Library of Babel, or just typed by monkeys, has, according to Shannon, the same information content as any Shakespearean tragedy and is indeed even more difficult to transmit, as Shakespeare's texts contain both sense and idiosyncrasies. An English text, like a text in any language, has many correlations and redundancies, like *u* following *q*, besides special features of the author's style, and therefore can, in principle, be compressed for more efficient transmission, whereas a nonsensical sequence of characters is irreducible, but contains no usable information.

It is amazing that some smart people engaged in lofty matters still do not understand how increasing entropy converts reversible fundamental laws into irreversible development. Stephen Hawking is praised for rescuing the loss of information in an object, say, a book, thrown into a black hole, by its imagined quantum evaporation, without suggesting in what sense this information would be restored. It would be easier to check this just by throwing a book into the fireplace. How would we restore the information from the ashes? Of course, restoring the text is not a problem, it is stored electronically and in many other copies – but where would the precious fine structure of the paper have gone?

The laureate of the 2004 Nobel Prize in Physics, Frank Wilczek (2016) says in a conversation with Edge: "The laws of physics had this property that seemed totally gratuitous, unnecessary to describe the world, in fact, kind of embarrassing. It's a famous problem called the "arrow of time". How can it be that the fundamental laws look the same forwards and backwards in time, and yet, the world doesn't?" He goes on to look for "deeper principles" in time asymmetries of obscure particle decays or something hidden in dark matter. Even if these oddities exist, they have nothing to do with the natural arrow of time that allows us to see why a movie wound backwards doesn't look right. Roger Penrose (2013) understands this: "In principle, one could also perform calculations in the reverse 'teleological' direction, because of the time-reversibility of the basic Newtonian laws. However, because of the second law of thermodynamics, whereby the entropy (or 'randomness') of a physical system increases with time in the natural world, such reverse-time calculations tend to be untrustworthy."

It is a drive to more probable states, from an intact to a broken glass, that defines the natural arrow of time. Highly improbable states, like life itself, do emerge but at the price of a growing entropy in the environment. Ultimately, we exist because the Sun is burning itself out. In his old age, Ilya Prigogine, laureate of the 1977 Nobel Prize in Chemistry, was also desperately looking for another arrow of time, although not as exotic as Wilczek's. I saw him presenting a talk in Paris: the whole situation was somewhat pathetic, to see him sitting on the stage, cared for by an assistant, and expounding his insights to a confused but attentive audience.

4.4 Chemical Bonds

Even in the days when physicists preferred the continuum, the atomism of John Dalton (1808) seemed a natural point of view to chemists. They operated with atomic and molecular weights, which were assumed to be fixed for each element and could be deduced from chemical experiments. Once the concept of chemical elements had been established, the next problem was to understand how they bond together in molecules. There were still many uncertainties, besides experimental errors. Two French chemists, Joseph Proust and Claude Louis Berthollet, disputed whether chemical compounds contain elements in definite proportions. It was indeed hard to decide. For example, some metal oxides have several forms, say, MO and MO_2, which may also be mixed, and in this case it had to be proven that it really was a mixture; besides, Berthollet, who thought the proportions to be variable, was a more authoritative figure, a member of every academy in sight, and the leading experimentalist who had established the chemical nomenclature together with Lavoisier. Moreover, it turned out later that there are in fact freak compounds that contain variable proportions of elements, just as Berthollet had insisted. Nevertheless, Proust, supported also by Jöns Jakob Berzelius, eventually won, and this gave a boost to the natural idea that molecules consist of atoms. Chemical species could now be assigned certain formulas, like H_2O for what had once been one of the fundamental elements: water.

But whether it was H_2O or HO or HO_2 or something else still had to be decided. Calculating the atomic weights of different elements depended on what proportions were assigned to the various compounds. John Dalton assumed that Nature obeys the principles of beauty and parsimony, cutting out irrelevant complications with the help of Occam's razor, whence the formula had to be HO. There was, however, a way to establish the proportion experimentally. In the same year of 1808, Joseph Louis Gay-Lussac published his law stating that the ratio between the volumes of reactant gases and gaseous products could be expressed in terms of simple whole numbers and found that water is formed by combining two volumes of hydrogen with one volume of oxygen. Further support for this law was given by Amedeo Avogadro

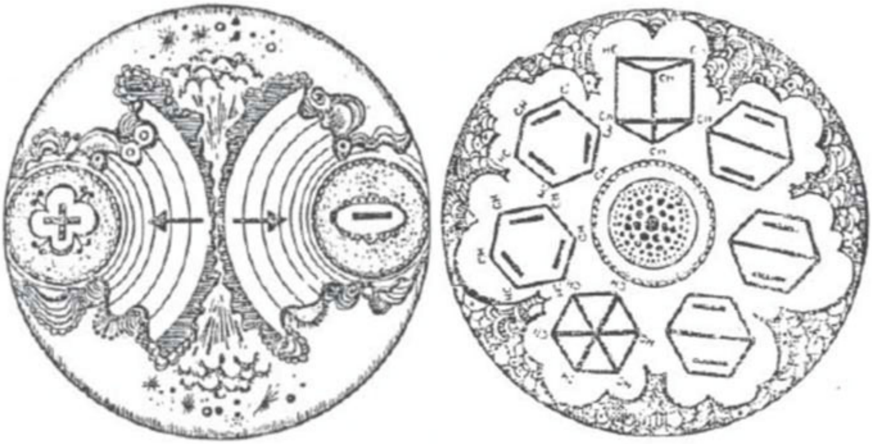


Fig. 4.4 Artist's rendering of chemical bonds. *Left*: An electrostatic bond between two oppositely charged parts. *Right*: Alternative covalent structures of benzene

(1811), stating that a given volume of any gas (we must now qualify, any ideal gas) contains the same number of molecules. Avogadro's work remained unnoticed for a while, but André-Marie Ampère, who was under the spotlight in Paris, came up with the same statement three years later.

There was, however, a mix-up between molecules and atoms. Gay-Lussac's result clearly indicated the formula we are accustomed to, but Dalton objected, pointing out that the reaction between nitrogen and oxygen, which was supposed to be written $N + O = NO$, should halve the gas volume, while in fact the volume remains the same. The true expression for this reaction is $N_2 + O_2 = 2NO$, which implies no volume change, but the notion that two identical atoms should combine appeared to be as unacceptable as gay marriage. The nature of chemical bonds was still totally unclear. In his *Opticks*, Isaac Newton (1704) mentioned atoms attached to each other by some force, without specifying which one; he would certainly not have thought it could be gravity, but what other force could he suggest? A century later, when electricity came into fashion, Humphry Davy and Jöns Jakob Berzelius advocated an electrostatic theory: molecules should combine positive and negative partners, like in salt crystals before they separate into ions when they dissolve (see Fig. 4.4). By this theory, the bond should be the strongest when the connected atoms were as different as possible. From this point of view, symmetric molecules were blatantly unnatural.

Electrostatic theory did not apply at all to the numerous compounds of carbon. Carbon atoms are capable of connecting in long chains decorated by other elements, primarily hydrogen, but also oxygen, nitrogen, and others. This leads to an infinite variety of structures, studied by a dedicated branch of chemistry. The splitting off of what came to be called organic chemistry was caused by the belief that organic compounds could only be produced by living organisms endowed with "vital force". The apparatus for their synthesis had to be complex and alive, while the chemist's coarse tools were only good for analysis, for breaking things up. This prejudice survived till the 19th century. The first synthesis of an organic compound (a constituent of urine) from inorganic salts had to wait until 1828, and it was still not quite convincing, since urine was just a waste product. Reference to the "vital force" was also convenient for explaining away the failure to understand the nature of chemical bonds in organic compounds: they are organic, after all, and are not obliged to obey the same rules as inorganic salts and oxides.

This could not go on for long. More carbon compounds were synthesized, most notably, aniline dyes, which acquired a high commercial value in the second half of the 19th century. This heated up the interest in organic compounds, and the "vital force" ceased to be a serious argument in the oncoming rational age. The basic notion governing the structure of molecules became *valency*, denoting the ability of an element to form a certain number of chemical bonds. The valency of an element can generally be deduced from its position in Mendeleev's table; this is what determines the periodicity in chemical properties. The underlying cause, anticipating the quantum-mechanical theory we will come to in Sect. 6.4, is the ability of an atom to supply either electrons or shared electron orbits to a *covalent* bond, and the valency will usually be the smallest of the two numbers. Carbon, and also silicon which sits just below carbon in the Periodic Table, has the highest valency, four, and this is responsible for the richness of carbon chemistry.

Recognition of the tetravalence of carbon by August Kekulé (1857) was a major step forward in uncovering the structure of organic compounds. Archibald Couper and Alexander Butlerov were his competitors in understanding the structure of the various molecules. The non-trivial observation was that carbon atoms can connect themselves by double and even triple bonds, as well as single ones, and this was what led, for example, to the difference in composition of ethane C_2H_6, ethylene C_2H_4, and acetylene C_2H_2. Understanding the ring structure of benzene C_6H_6 was a major riddle; the solution came to Kekulé (1865) as he slumbered on a London double-decker, in a dream about Ouroboros, a serpent eating its tail (Fig. 4.5). The ultimate understanding of the nature of chemical bonds came only with the advent of quantum mechanics (Sect. 6.4).

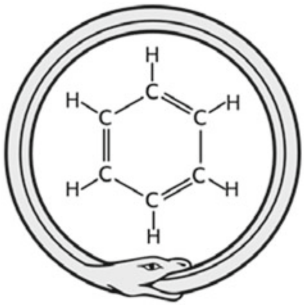


Fig. 4.5 Ouroboros, Kekulé's inspiration

Starting from the late 19th century, organic chemistry became the basis of major industries, as synthetic dyes were followed by various drugs and polymers; the processing of coal and oil supplied raw materials. It also grew into a kind of formal structure, with a well-organized nomenclature that allowed anyone to reconstitute the structural formula from the name of a compound. There is even a convention to place the prefixes referring to smaller groups first, e.g., placing methyl CH_3 before ethyl C_2H_5. A late-19th century Russian chemist formulated this rule as "children and dogs go ahead". While, with my natural clumsiness, I hated lab work, as an undergraduate, I enjoyed formal exercises in drawing a sequence of reactions leading from one organic substance to another, something that often had almost mathematical clarity.

4.5 Chemical Kinetics

All chemical compounds handled by the 19th century chemists were, by definition, stable. One needs modern sophisticated devices to capture evanescent creatures expiring in nanoseconds. The question is how do they transmute into one another. Chemical reactions are habitually expressed by stoichiometric equations, like $O_2 + 2H_2 = 2H_2O$. A natural but wrong idea is a naive "mass action" law: the rate of a chemical reaction is proportional to concentrations of reactants, taken in powers corresponding to their stoichiometric coefficients. Does this mean that the rate of this reaction should be proportional to the concentration of oxygen times the concentration of hydrogen squared? In no way! Triple collisions of molecules are rare, and the actual reaction mechanism does not follow its stoichiometric equation. The mass action law may only be applied to elementary intermediate steps, which are rarely evident.

A stable compound needs a substantial perturbation to enter a chemical reaction. The initial and finite configurations of atoms or atomic groups can be imagined as two isolated valleys (Fig. 4.6) surrounded by high mountain ranges and connected by a mountain pass. The landscape in the picture is an energy landscape, and the pass is a saddle point of the energy levels. The elevation of the saddle point is not related to the elevation of the valleys, and this barrier has to be overcome in order to pass to a friendlier and greener valley. This hike requires an effort, though

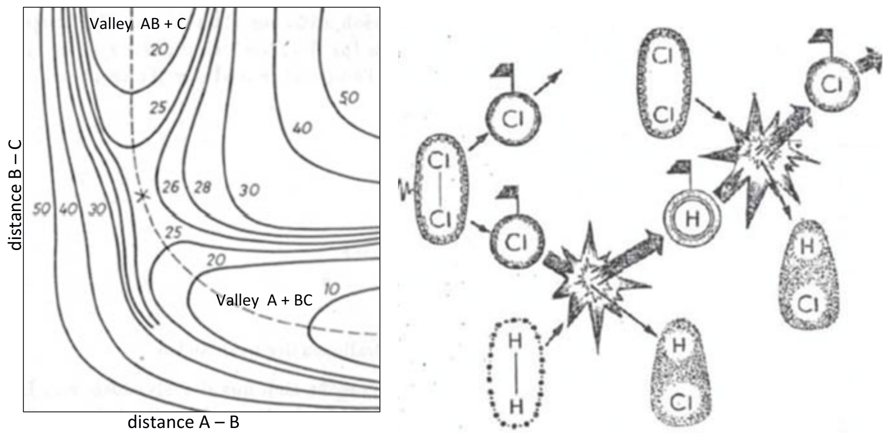


Fig. 4.6 *Left*: Energy landscape of a reaction A + BC = AB + C (the numbers are arbitrary). *Right*: Artist's rendering of a chain reaction involving hydrogen and chlorine radicals

lower than climbing steep ridges. Overcoming the energy barrier is helped by the Maxwell distribution: energetic molecules, always present, even if in small fractions, will manage to go through. The probability of passing is, in this scenario, an exponentially decreasing function of the ratio of the height of the energy barrier to the thermal energy. This dependence is captured by the temperature dependence of the reaction rate proposed by Svante Arrhenius (1889). His inspiration very likely went back, not all the way to Maxwell, but to Jacobus van 't Hoff, the first winner of the Nobel Prize in Chemistry, who established a similar formula for chemical equilibrium guiding the distribution of molecules shuttling between two valleys, dependent on the ratio of their elevations. All this soon became the staple of undergraduate courses.

Activation energies may be rather high when the reaction involves breaking strong covalent bonds; accordingly, the temperature dependence is steep, and the reactions go slowly at moderate temperatures that are especially favorable for the synthesis of delicate and vulnerable products. Fire is one of many reactions that is strongly influenced by temperature: wood does not burn at all (and neither do manuscripts) before being ignited, but then it burns all the way leaving only ash. The most common way to carry out chemical reactions in a subtle and controlled way is *catalysis*. A catalyst can be compared to a mountain guide leading hikers through caves and crevices and bypassing the high elevation of the main road. Catalysts, while taking part in intermediate reaction stages, are not consumed in the finished process. Some catalysts are dissolved in relatively small amounts in the reaction mixture, but most effective industrial catalysts are solid. Reactants are adsorbed on the catalytic surface. This softens their structure, they react in a desired direction, and the products desorb. Designing efficient catalysts is more of an art than a science. Catalytic processes are subtle and all catalysts eventually lose their activity; they "get tired". This may not even be caused by a change in their chemistry, but just by a subtle alteration in their surface texture.

Modern tools, from electron microscopy to atomic force microscopy (AFM), are able to detect minute details of catalytic structure and the detailed course of a reaction (Fig. 4.7, left). Gerhard Ertl won the 2007 Nobel Prize in Chemistry for his

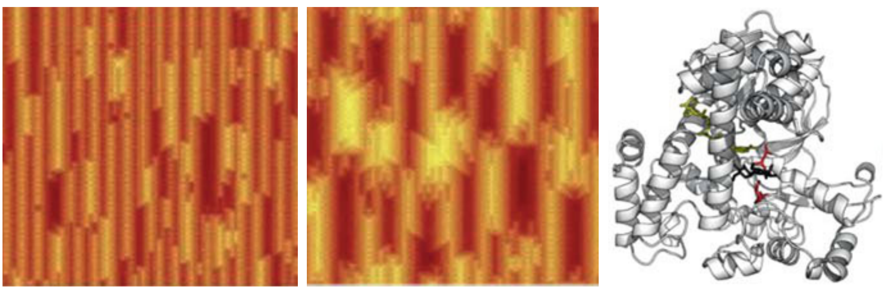


Fig. 4.7 Simulated restructuring of a catalytic surface when the reactant concentration changes from a low level (*left*) to a higher one (*center*) (Monine et al, 2004). *Right*: Model of an enzyme. The substrate molecule shown in *black* is held by active sites

study of fine nanoscale patterns on a catalytic surface [Ertl (1991)]. AFM studies reveal the atomic detail of a catalytic surface. Studies of this kind are carried out in high vacuum on a particular face of a catalytic crystal (usually platinum), and relate to practical problems about as much as extragalactic astronomy relates to aerial navigation.

Nature is masterful in handling complex chemical reactions catalyzed by enzymes – huge protein molecules with a structure finely adjusted to capture a certain chemical species ("substrate"), carefully guide it through a sequence of required transformations, and release the product (Fig. 4.7, right). Enzymes are highly specialized and are often able to carry out only a single function. They are highly flexible, changing their conformation in the course of the process. They are themselves synthesized by the intracellular machinery when their service is required, and disintegrate when they are no longer needed. Industrial catalysts are not as specialized, and are as unlike enzymes as computers are unlike the brain.

When catalysts are not involved, a common mechanism is a chain reaction. Walther Nernst (1918), taking note of an excessively high yield per photon in a photo-induced reaction between hydrogen and chlorine, suggested that the reaction propagates through a chain of unstable intermediates (Fig. 4.6). These intermediates are *radicals*, that is, they have an unsaturated bond, or in quantum terms, an orbit with a single electron, so once they have formed, they react eagerly, producing one or more new radicals alongside the desired reaction product. The quantitative theory of chain reactions was developed by Nikolay Semenov (1935); it brought him the 1956 Nobel Prize in Chemistry, shared with Cyril Hinshelwood, who recognized adsorption as the rate-limiting step in catalysis.

Clearly, if more than one radical is produced and radicals do not recombine or escape fast enough, their numbers will grow exponentially and the reaction will become explosive. Chain reactions soon became significant on a deeper and more dangerous level than in chemistry. Nuclear fission was discovered just in time for World War II, and there were no authorities who understood its military significance well enough to suppress the news. A neutron triggering the fission of a heavy atom produces more neutrons, and the resulting release of energy exceeded everything known prior to that. Suddenly, various people in the USA, Britain, and the USSR came along with theories of a strong explosion being caused by a point source. As this topic was immediately classified, there were priority disputes when the war was over. Certainly, everybody was right in this case, especially as this theory was easier than describing an explosion of the same intensity caused by the required amount of ordinary explosives.

More chemistry than physics was involved in the forthcoming work on nuclear weapons, which was largely delayed by a difficult process of separation of uranium isotopes. Theorists worked with simple models cracked numerically by female "computers", as they were called before the advent of what we now call computers. As John Wheeler (1998) recalls, while discussing nuclear fission with Bohr, he used the same language (a hiker in the mountains) for a nucleus deforming to overcome a potential barrier of nuclear forces before splitting. They asked for the advice of Eugene Wigner (1963 Nobel Prize in Physics), who was trained as a chemical engineer before diving into mathematical physics.

Further lessons in activation energy were learned in the development of the H-bomb, which needs the heat supplied by a fission bomb to start the nuclear fusion. Awful as this unworldly weapon is, destroying entire islands and filling the atmosphere with radioactive exhaust before all but underground tests were banned, Wheeler (1998) maintains in his memoirs that the hydrogen bomb kept the Cold War from getting hot and praises Edward Teller, the driving force behind the US H-bomb. Teller was awarded the first Ig Nobel Peace Prize for his efforts, while his Soviet counterpart Andrei Sakharov won the real Nobel Peace Prize – but not for his leading role in the H-bomb project.

Chapter 5
Condensed Matter

5.1 Crystals

The development of statistical physics in the 20th century was dominated by the struggle to understand intermolecular interactions in condensed media. Theory could advance from two opposite directions. One extreme was the total disorder of the ideal gas, which could be perturbed by adding weak interactions, as van der Waals did (Sect. 4.2). The other was the total order of a crystalline solid.

Order is more complicated than disorder: disorder is unique, while there are different kinds of order. Order is, however, rational, and had already been understood precisely in the 19th century. A crystal structure is determined by the structure of a unit cell, infinitely repeated by translation in three dimensions. There are 32 possible crystal classes. Johann Hessel (1831) derived them, without the benefit of mathematical group theory, by studying actual crystal forms. This classification, published in an obscure Physics Dictionary, went unnoticed until reprinted posthumously in 1897. Auguste Bravais (1850) proved the existence of 14 Bravais lattices, based on 14 types of unit cell that could be repeated by translating in all three dimensions to

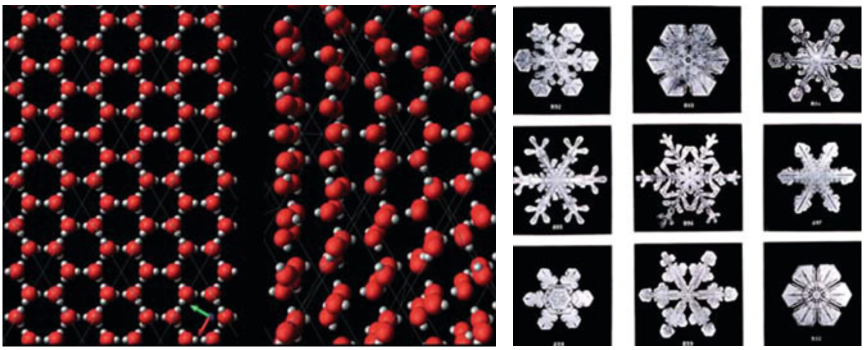


Fig. 5.1 *Left*: Two views of the crystalline structure of ice. *Right*: Snowflakes

L. Pismen, *The Swings of Science*,
https://doi.org/10.1007/978-3-319-99777-3_5

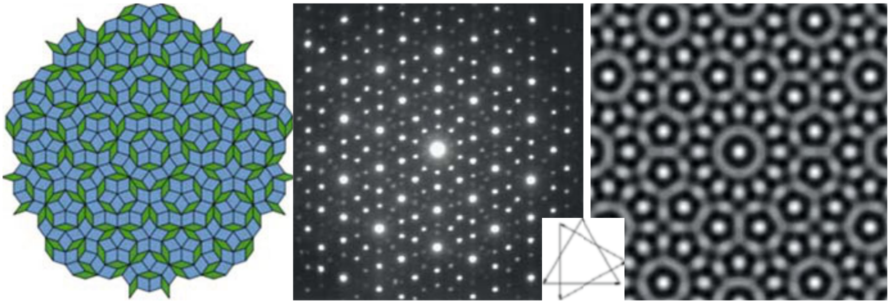


Fig. 5.2 *Left*: The Penrose tiling. *Center*: Electron diffraction pattern of an icosahedral quasicrystal. *Right*: A quasicrystalline structure obtained as a superposition of six waves, directed as shown in the *inset*

build up all crystalline structures. Additional crystal classes can be obtained by rotating unit cells. The crystalline structure is revealed by the shape of slowly growing crystals, as we see in the shapes of snowflakes, which retain the hexagonal symmetry of the crystalline lattice (Fig. 5.1).

Crystals may possess an n-fold rotation axis with n equal to 2, 3, 4, or 6. When Dan Shechtman (1984) found the forbidden 5-fold symmetry in a diffraction pattern, he hesitated to publish. His hesitation was unwarranted. Quasiperiodic structures were predicted by Harald Bohr[1] as projections of regular crystals in higher dimensions (Bohr, 1925), and in the 1970s Roger Penrose invented a quasiperiodic tiling of a plane bearing his name. This pattern never repeats itself, but has a simple structure built up of just two types of rhombic tiles. A quasiperiodic structure in the plane obtained by superposition of six waves looks similar to the diffraction pattern of a quasicrystalline material (Fig. 5.2). The story came to a happy end when Shechtman was awarded the 2011 Nobel Prize in Chemistry.

Before the advent of quantum mechanics, crystal structures were mainly investigated in relation with the optical, magnetic, and thermal properties of solids. Mechanical properties are not directly related to crystalline structure, but largely depend on the nature of the bonds between atoms or molecules in the crystal maze.

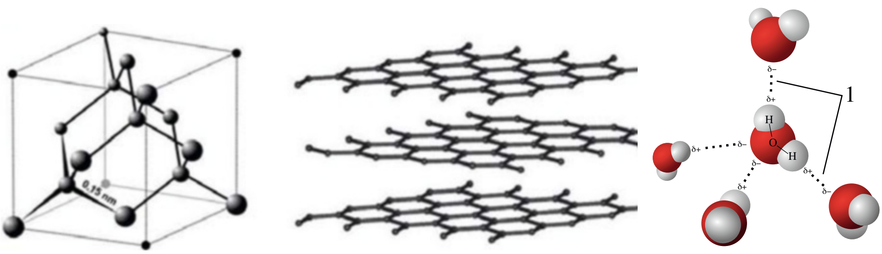


Fig. 5.3 The structure of diamond (*left*) and graphite (*center*). *Right*: The hydrogen bond in ice

[1] The mathematician brother of Niels Bohr.

The strongest bonds are covalent and ionic (Sect. 4.4). Metal atoms are bound by electrons delocalized throughout the crystal, and this also binds them strongly. A weaker bond exists between dipoles, molecules with a predominant positive charge at one end and negative charge at the other; an example is the hydrogen bond connecting water molecules in ice. The weakest are van der Waals bonds between induced dipoles emerging due to correlated fluctuations of electron distribution in nearby molecules. Diamond owes its strength to covalent chemical bonds between its atoms. Within graphite layers, chemical bonds are similar to those in a benzene ring, already somewhat weaker, but layers are bound only by van der Waals bonds, the weakest of all (Fig. 5.3).

Most solids, except carefully grown crystals, are polycrystalline, and this usually makes them much weaker mechanically[2]. Even single crystals are weakened by *dislocations*. The mathematical theory of dislocations was initiated by Vito Volterra (1907) in the context of continuous media. Jan Burgers (1940) described crystal dislocations in terms of the failure of a contour to close when it is followed along the edges of a fixed number of unit cells. The distance and direction between the start and end points is defined by the Burgers vector; it is parallel to the contour for an *edge* dislocation, and perpendicular for a *screw* dislocation (Fig. 5.4). The theory of crystal defects has developed into an important branch of differential geometry (Kröner, 1966). Its applications extend to relativity theory. In Sect. 1.4, I mentioned the extension of Einstein's general relativity including torsion (Hehl et al, 1976). A path in the universe with torsion fails to close, just as it does not close when it goes around a crystal dislocation.

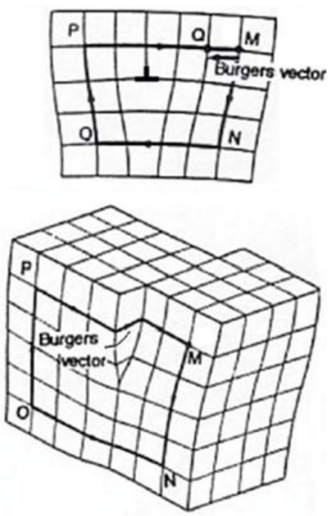


Fig. 5.4 An edge dislocation (*upper*) and a screw dislocation (*lower*)

5.2 Phase Transitions

Imperfections of crystalline order are not necessarily static. Atoms or molecules vibrate around their equilibrium positions. Studies of lattice vibrations were initially motivated by the problem of calculating the heat capacities of solids. Peter Debye, whom we met in Sect. 3.4, and Einstein himself were involved in this problem, but the major advance was made by Born and von Kármán (1912) with their theory of

[2] Weakening can be reversed by clever structuring. In a talk at the 2018 APS March meeting, Pupa Gilbert attributed the extraordinary strength of dental enamel to the misalignment of the microscopic crystals that prevents cracks from propagating.

collective vibrations in lattices. Like all waves, such vibrations also have the properties of particles in quantum theory, and when quantized, they are called quasiparticles – *phonons*. The name comes from the relation to sound, but in classical theory (Sect. 2.5), waves are periodic deformations of a continuous medium rather than of a lattice, and only longitudinal waves change the local density and thus get classified as sound waves. Max Born went on to become one of the major figures in the development of quantum mechanics, and this brought him the 1954 Nobel Prize in Physics, while Theodor von Kármán concentrated on elasticity and hydrodynamics.

Yakov Frenkel (1946) called attention to the dynamic imperfections of lattices that lead to the formation of vacant sites (holes) and atoms or ions drifting in the interstices of the crystalline grid. Imperfections appear, are healed, appear again at different locations, and "evaporate" upon reaching the surface. He also drew a picture of "nomadic" and "settled" molecules in fluids determining the fluid viscosity and diffusivity. In this way, the difference between solids and liquids is just quantitative: as vibrations strengthen and holes and nomadic atoms multiply, order is lost and the solid melts. However, this does not happen in a continuous way. Melting and solidification are discontinuous, *first-order* transitions, and liquid water can still exist below the melting point in a supercooled state (Fig. 5.5 left).

Frenkel tried hard to explain this discontinuity in his semi-quantitative way, and not very persuasively. He was writing his classical book during the war years, lacking an office as well as living space, after his Physico-Technical Institute was evacuated from Leningrad – but scientists were taken care of in the end, otherwise he would have died of hunger in the besieged city, as hundreds of thousands did. There was no humanitarian aid from any side in those cruel times. Frenkel best characterizes his approach himself (Frenkel, 1946a): "The more complex the system under consideration, the simpler, by necessity, should be its theoretical description. [...] The theorist is similar in this respect to a cartoonist who, unlike a photographer, does not need to reproduce the subject in every detail, but has to simplify and schematize it to identify and emphasize its characteristic features. Photographic accuracy is possible, and should be required, only for the description of the simplest systems. The theory of complex systems should be only a good 'caricature' of these systems, exaggerating those of their properties which are the most typical, and deliberately ignoring all the other – inessential – properties."

Frenkel's approach has not been widely imitated. A much more complicated but less suggestive theory of liquids and dense gases was presented in the early postwar years by the Bogoliubov–Born–Green–Kirkwood–Yvon (BBGKY) hierarchy, which included correlations between multiple molecules in the course of multiple collisions. As in all theories of this kind, realistic approximations require a closure of the hierarchy at some stage. Modern approaches rarely follow Frenkel's advice. On the one hand, nice simple qualitative explanations of complex phenomena have already been invented, and the further we go, the more complex are those that remain unexplained. On the other hand, computers bring more opportunities to crack complex problems by brute force, or at least crack their outer shell.

The BBGKY hierarchy could not explain an abrupt phase transition either, but an utterly simple theory was already in place. Lev Landau (1937) cut the Gordian

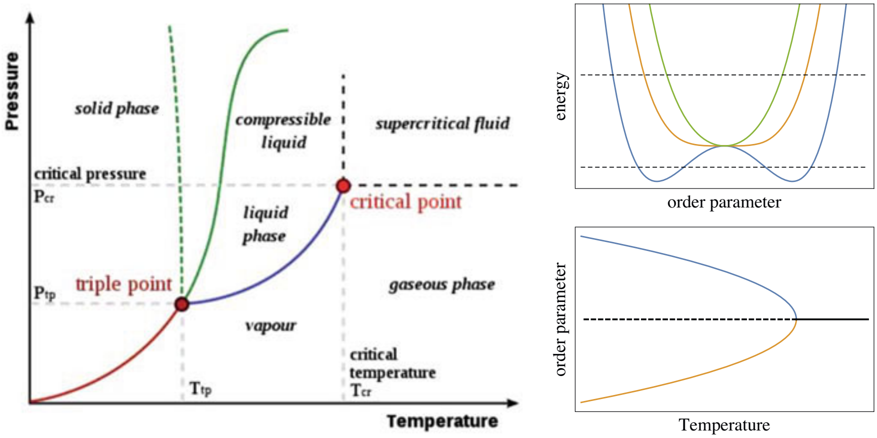


Fig. 5.5 *Left*: A typical phase diagram. The dashed line going up from the triple point bounds the area where a supercooled liquid exists. *Top right*: Different shapes of the Landau energy as a function of the order parameter. The *middle curve* corresponds to the critical temperature, and two minima appear as the energy is lowered further. *Bottom right*: The order parameter as a function of temperature. Two stable solutions exist below the critical point

knot by abstracting from all particulars and introducing a single characteristic: the *order parameter* that distinguishes one state from the other. The equations defining the value of the order parameter are utterly simple: it is enough to express the free energy as a function of the order parameter in such a way that it may have two minima under some conditions. A quartic polynomial will suffice, with its coefficients, depending on some control variable, e.g., temperature, changing in such a way that the number of solutions changes when a critical point is passed. Near a critical point, the dependence of the energy on the order parameter and the dependence of the order parameter on the control variable should look like in Fig. 5.5 (right). If the order parameter is identified with the density, the difference between its values in the two phases is proportional to the square root of the deviation from the critical point, exactly as in van der Waals' theory (Sect. 4.2). Further away, these curves will be distorted, but not in a qualitative way. Simplicity and universality, independent of whatever physics is involved, explain the outstanding popularity of Landau's approach, applied to any kind of phase transition. The choice of term itself is unfortunate: it is really not a parameter but a *field*, as it may change in space and time, and the Landau energy is usually supplemented by simplest possible terms depending on spatial derivatives of the order parameter. An extension of Landau's theory is the Cahn–Hilliard equation, which derives from Landau's energy, not just the rate of change of the order parameter with time, but also its flux, and thereby takes into account overall conservation of material (Cahn and Hilliard, 1958).

The knot was, however, not yet quite untied, as neither Landau's theory nor its modification took into account one crucial factor – fluctuations, which are the strongest exactly where, formally, Landau's approach works best – near a critical point. The phase transition between liquid and gas at the critical point does not involve any jump, it is a *second-order* transition. It has long been known that a fluid loses transparency close to the liquid–gas critical point. This *critical opalescence* was attributed to the scattering of light by density fluctuations. Leonard Ornstein and Frederik Zernike showed that the density correlation length, an estimate of the size of fluctuations, grows in inverse proportion to the square root of the deviation from the critical point, so that it diverges there (Ornstein and Zernike, 1914). Sufficiently close to criticality, the size of fluctuations becomes about the same as the wavelength of light, and therefore the fluid becomes opaque.

Experiment showed that the jump between the gas and liquid densities for different fluids is proportional to the cubic rather than the square root of the deviation from the critical point. The breakthrough in understanding near-critical phenomena came with the renormalization group theory of Kenneth Wilson (1971). Renormalization involves a repeated increase in scale that goes in parallel with the increase in the size of fluctuations as the critical point is approached. These transformations cause the system parameters to change, and if a critical point exists, they will come to a *fixed point*, where it becomes possible to read off things like the power dependence of the various properties on the deviation from the critical point. Wilson was almost fired from Cornell, as he did not publish anything while working on his theory, but was rewarded by the 1982 Nobel Prize in Physics. The renormalization group is similar to the techniques used in particle physics to eliminate divergencies. Shortly after inventing this technique, Wilson himself applied it to explain the confinement of quarks in hadrons.

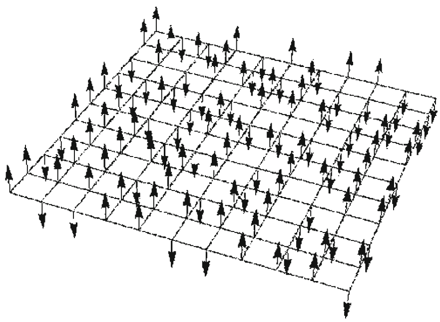


Fig. 5.6 The Ising model

Another thoroughly studied phase transition, motivated by magnetization, involved a toy model invented by Ernst Ising (1925). There are two opposite orientations of tiny magnets placed on the nodes of a two-dimensional grid. It is a square grid in Fig. 5.6, but it could be a different one. The magnets interact with their closest neighbors, and switch their direction if it lowers their energy. Ising himself did not think that a phase transition would be possible in this system, but a phase transition does indeed take place, and as the interaction strength grows, all the little magnets orient in the same way, so that the grid is "magnetized". The toy model became important when Lars Onsager (1944) solved it precisely. It was a very long and complicated proof, which was later shortened and clarified. The analytic solution made it possible to understand in detail the behavior near the critical point, and this could later be compared with what was predicted by renormalization group theory.

5.3 Surfaces

Sydney Goldstein (1969) observed that "in the 1920s we were complaining that it was impossible to remember while reading Lamb (1895) that water is wet". We start feeling this when we look closely at surfaces separating liquids from gases or solids. This lay outside the scope of classical hydrodynamics, falling into the domain of molecular physics and physical chemistry. Molecules within the bulk of a liquid interact on all sides with molecules of the same kind, while molecules near the surface have agreeable neighbors only on one side (Fig. 5.7 left). If this is the boundary with the vapor phase, there are few molecules of the same kind there. If this is the boundary with another fluid phase, interactions with molecules on the other side are less favorable – this is why the phases are separated.

In both cases, the energy of the surface molecules is higher, and there is a net force acting to reduce the extent of the surface at a given volume. Therefore when a flat surface is perturbed, convex patches are pressed downward and concave patches upward restoring the flat shape. The surface can support a light weight compensated by the upward pressure of concave troughs; a water strider uses this effect (Fig. 5.7 center). Jesus could walk on the waters of Lake of Galilee as well, if strongly levitated by his Father or moving faster than the viscous relaxation time, so that the shape of the surface would be too tardy to respond to his steps.

For the same reason, droplets suspended in air or in another fluid acquire a spherical shape to reduce their energy; this shape can only be distorted by gravity or flow in the surrounding medium. The same effect causes a liquid cylinder to undulate and eventually break into droplets. Remarkably, this Rayleigh–Plateau instability saturates due to weak elasticity in soft solids (Mora et al, 2010), relaxing to a stable undulating shape, as in Fig. 5.7 (bottom left). The higher energy of convex areas also enhances evaporation or slow dissolution in a surrounding fluid. This is the Kelvin effect, named in honor of William Thomson (1871). It causes the *ripening*, or coars-

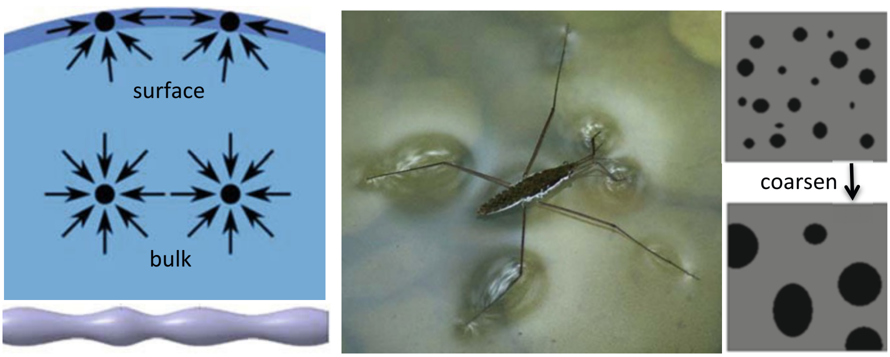


Fig. 5.7 Surface tension. *Top left*: Forces acting on a molecule in the bulk of a fluid and on its surface. *Bottom left*: Perturbation of a liquid jet or a soft solid cylinder. *Center*: Surface tension sustains a water strider on the surface. *Right*: Matthew's law for droplets

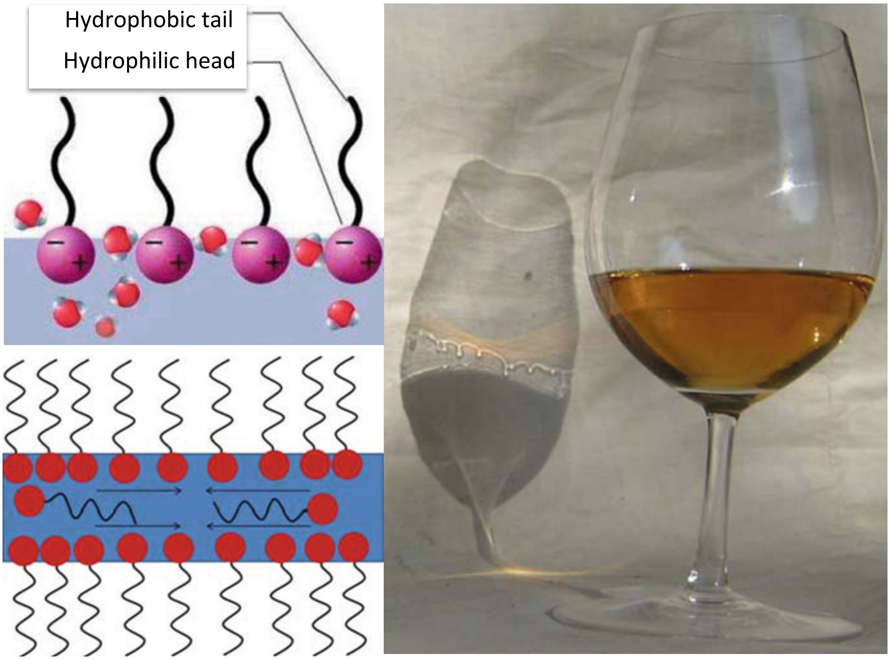


Fig. 5.8 *Top left*: A surfactant on a water surface. *Bottom left*: A soap film. *Arrows* indicate the direction of the Marangoni force. *Right*: Wine tears (seen in the shadow)

ening, phenomenon: smaller droplets dissolve and larger ones grow at their expense (Fig. 5.7 right). This is one of the realizations of the evangelical Matthew's law: "For whosoever hath, to him shall be given, and he shall have more abundance: but whosoever hath not, from him shall be taken away even that he hath" (King James Bible). If we wait sufficiently long, only a single drop will remain. Societies prevent Matthew's law taking its course to the end, either by taxation and social security or by spontaneous turmoil. It is often also advantageous to avoid it in two-phase fluids. For this purpose, *surfactants* are used: amphiphilic molecules tending to congregate on the surface, say, with the hydrophilic head in water and the hydrophobic tail in air or oil (Fig. 5.8 left); this reduces surface tension and favors a surface area large enough to accommodate the surfactant. Soap is the most familiar surfactant, stabilizing thin water films in froth or soap bubbles.

The surface energy can also be reduced without changing the surface area if some patch of the surface has a lower surface tension because of a higher local temperature or concentration of some admixture with a lower surface tension. The attention of Lord Kelvin's brother James Thomson (1855) was caught by "wine tears" (Fig. 5.8 right). He was certainly not the first to see them, but he was apparently the first to publish a qualitative explanation. Alcohol has a lower surface tension than water and it also evaporates faster than water because its boiling point is lower. Therefore a wetting film on the glass becomes bereft of alcohol, its surface tension grows, and wine flows along its surface to replenish alcohol and reduce the surface energy. For the same reason, a uniform surfactant concentration is restored in a soap film (Fig. 5.8 bottom left). This effect bears the name of Carlo Marangoni (1869), who explored it in a dissertation that long went unnoticed, however. Skip Scriven who thoroughly studied and popularized the Marangoni effect in the 1960s called him "an obscure Italian physicist whom I made famous". He is famous now indeed, at least in a restricted circle; there is even an International Marangoni Association which organizes regular conferences, an honor rarely if ever bestowed on anyone.

Enlightened lay readers might have been persuaded by popular science publications that quantum mechanics quantizes everything on sufficiently small scales. We are sometimes able, however, to move far into the 20th and 21st centuries and to the scale of nanometers without feeling quantum effects. Van der Waals extended his law to devise an equation describing coexistent liquid and vapor phases, together with their boundary (van der Waals, 1894). This boundary is only a few nanometers thick but can be resolved by modern methods. It permanently fluctuates, and a continuum description, which is a version of the ubiquitous Landau model (Sect. 5.2), describes it in a very rough way, but makes it possible to estimate the surface tension using the intermolecular interaction energies.

The analogy with liquid droplets goes to still shorter scales. When nuclear fission was discovered, there was an immediate drive to explain it, even in a qualitative way. Both Yakov Frenkel (1939) and independently Niels Bohr and John Wheeler (1939) used the analogy with a droplet that deforms and may eventually break when disturbed by an external force. A heavy uranium nucleus behaves in the same way under the action of impinging neutrons. It is kept compact by strong nuclear forces but deforms "from orange to cucumber to cashew nut", as Wheeler (1998) puts it, and then – bang! – breaks into the more stable nuclei of elements in the middle of Mendeleev's table.

5.4 Three-Phase Lines

The situation becomes more interesting and even controversial when solids are involved. A sessile droplet forms a contact angle depending on the strength of interactions between the liquid and the solid. If these interactions are favorable, the angle becomes smaller, and eventually the liquid wets the solid surface completely. A fluid rises up or moves down a capillary, respectively, when the contact angle is acute or obtuse (Fig. 5.9). The contact angle is defined by the classical Young–Laplace formula (Young, 1805)[3], which, however, contains unmeasurable surface tensions on the solid interface. The three-phase boundary is the *contact line*. A better but less straightforward explanation involves molecular interactions between the solid and the liquid which act either to repel the liquid surface from the solid by "disjoining pressure" or attract it when this pressure becomes negative (Derjaguin et al, 1987).

[3] Thomas Young actually preferred to explain relations by words rather than writing them in a mathematical form.

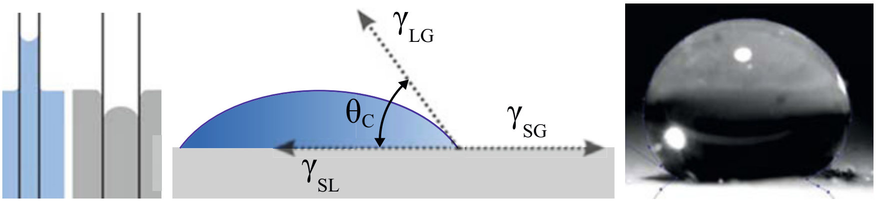


Fig. 5.9 *Left*: A wetting and a non-wetting liquid in a capillary. *Center*: A sessile droplet, with *arrows* showing forces due to gas–liquid (γ_{LG}), solid–liquid (γ_{SL}), and solid–gas (γ_{SG}) surface tensions. *Right*: A droplet on a non-wetted lotus leaf

In this description, going back to pre-war work by Boris Derjaguin, the surface is bent close to the contact line – usually very close, in a nanometer range.

When a droplet starts moving, as on a windscreen in rain, the advancing angle increases and the receding one decreases. When a plate is pulled out of a wetting liquid, it carries a wetting film against gravity – the effect bears great names (Levich and Landau, 1942) and is still being studied right down to minor details. Experiment confirms the changes in the dynamic contact angles, though precise data are hard to come by because they depend on surface roughness, even on a very fine scale. But mathematically, there is a severe problem, even on an ideally smooth surface.

The Clay Mathematics Institute has called the existence of smooth solutions of the Navier–Stokes equation one of the seven most important open problems in mathematics and has offered a million dollar prize for a solution or a counterexample. It is a purely mathematical problem, and the prize is likely to remain unclaimed till inflation reduces its value to the price of a quick lunch. No physicist would imagine that viscosity does not suffice to regularize solutions in the bulk. But when the fluid has free boundaries, it's a completely different matter. If you try to solve the moving contact line problem in a straightforward manner –something which, however, can be done analytically only at the cost of some simplifying assumptions such as neglecting surface deformation, which is, of course, wrong – you come to a paradox: an infinite force is needed to move the contact line (Huh and Scriven, 1971)! This is a good indication that the solution does not exist in the framework of classical hydrodynamics – I would redirect the million dollar prize to encourage mathematicians to prove it.

There are many ways to resolve this singularity (Bonn et al, 2009): through hydrodynamic slip at the solid surface, the formation of an ultra-thin precursor layer stretching on the solid ahead or behind the contact line, or the diffuse character of the interface. All of them, when matched to macroscopic flow, yield qualitatively similar results, making the advance velocity inversely proportional to the logarithm of the ratio of a macroscopic scale (e.g., the size of a droplet) to the microscopic scale, whatever it is, and going down to zero when the microscopic scale vanishes, as it does in the classical formulation.

This is a bad news for engineers modeling such important processes as coating or immersion lithography. Their common numerical methods generally do not fare

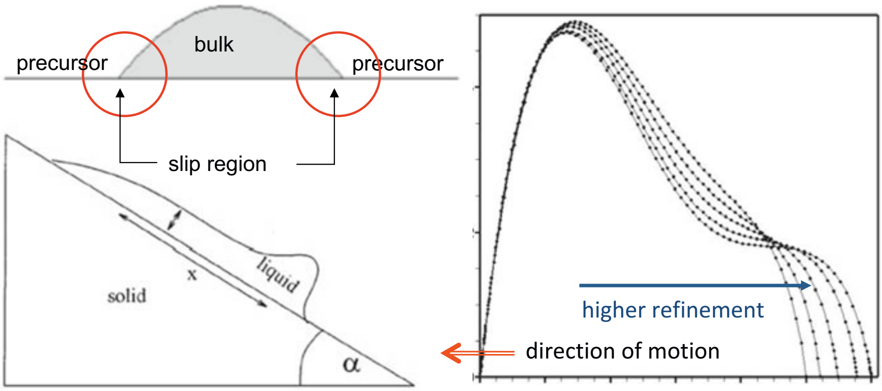


Fig. 5.10 *Top left*: A sessile droplet and schematically indicated locations of a precursor layer and hydrodynamic slip. *Bottom left*: A large droplet on an inclined plane, shaped like a “boa that swallowed an elephant”. *Right*: The shape changes with refined grid near the contact line

well when moving interfaces are concerned, but when a singularity is encountered, as in the moving contact line or in the break-up of a jet into droplets, the result becomes critically dependent on the size of the discretizing grid. Computers are still not powerful enough to refine the grid to near molecular fineness, and their programmers are not ready to include molecular interactions prevailing at these scales. The velocity of a sliding droplet and its shape depend on the grid size near the contact line, and the dependence never saturates until molecular scales are reached (see Fig. 5.10). This inconvenient fact is commonly swept under a rug.

5.5 Liquid Crystals

When a solid melts, its crystalline order is lost, but some liquids retain vestiges of order. These are *liquid crystals*. I will not go into a long history of different kinds of liquid crystal, starting from an accidental discovery by a botanist Friedrich Reinitzer (1888). The most comprehensive modern treatise is the book by de Gennes and Prost (1993). As already mentioned in Sect. 5.1, there are many kinds of order, while disorder is unique; the principal liquid crystal phases are sketched in Fig. 5.11. The most common one is nematic; it retains orientational order and is otherwise disordered. In the cholesteric phase, the prevailing orientation rotates around some axis. In smectics, molecules are layered; within each layer, they are oriented normally in the A phase, at a certain angle in the C phase, and at an angle rotated between one layer and the next in the C* phase; within each layer, their positions are disordered. The schematic pictures should not be taken literally. The spindles just show the orientation of molecules, characterized by a *director*, which is like a vector without an arrow. The preferred order is determined by interactions between molecules

Fig. 5.11 Schemes of liquid crystalline phases. From *left* to *right*: nematic, smectic A, smectic C, cholesteric, smectic C*

that depend on their chemical structure and shape. The molecules are elongated and usually contain some rigid elements, like double bonds and benzene rings. As in all phase transitions, order diminishes as temperature grows, so that, for example, a smectic can "melt" into a nematic and then to an isotropic liquid.

From the practical point of view, what is special in liquid crystals is their optical properties. Due to their anisotropy, liquid crystals polarize light. As a transverse wave, light has two polarization directions: a liquid crystal transmits the one aligned with its director, and also rotates the polarization. The orientation of the director can be easily controlled by treating the confining surfaces and by using electric or magnetic fields. Our computer and TV screens are built up of liquid crystal pixels, as shown schematically in Fig. 5.12. The light beam passes through a polarizer which polarizes it parallel to the orientation of the liquid crystal set by treating its surface. The orientation of both the polarizer and the liquid crystal at the other end are rotated through 90°, and light, rotated by the twisted liquid crystal, passes through (left panel). When the voltage is turned on, the liquid crystal orients itself along the beam direction, the polarization of light is not rotated any more, and the beam is blocked by the adverse polarizer at the exit (center panel of Fig. 5.12).

Liquid crystals, like solids, possess elasticity that tends to orient them uniformly. Compared to solid crystals, elasticity is limited. In nematics, it affects only the change of orientation, which can be of three kinds: splay, twist, and bend, as illustrated in the right panel of Fig. 5.12. Since there is no translational order, nematics can flow like normal fluids, only their viscosity is anisotropic. Smectics are already sensitive to the distortion of the shape of their layers and are only mobile within the layers. Variegated liquid crystal textures containing *defects* are what is most interesting from a scientific point of view. If an isotropic liquid is "frozen" into the nematic state, the orientation will be different at different locations and a lot of defects will appear. These defects will gradually collide and annihilate, and the texture will become more regular, but in many cases, not all defects will disappear.

A defect is not just a flaw or lack of something: it is a very precise mathematical notion that belongs to the field of *topology*. This involves sophisticated theory concerned with the properties of space that are preserved under continuous deformations and can only change when, for example, holes are torn. Topology plays an

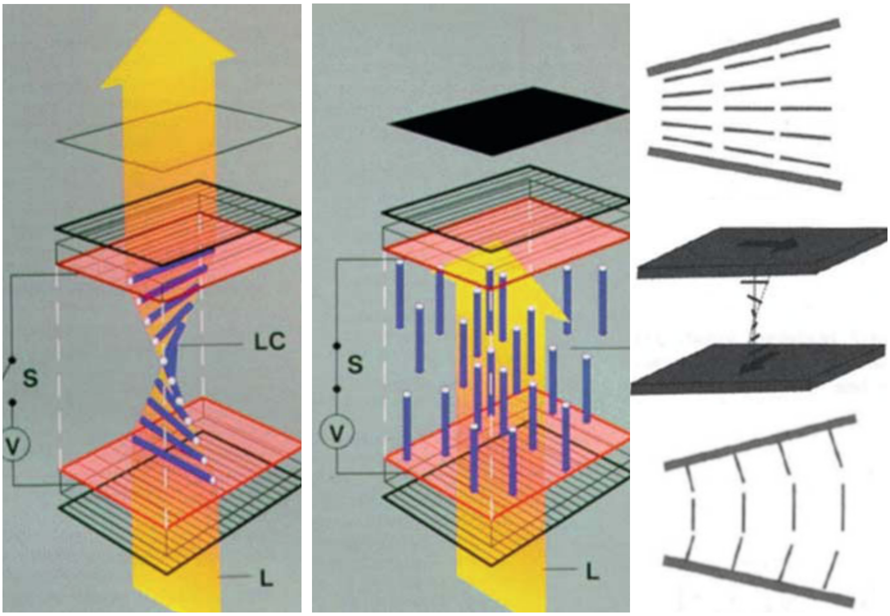


Fig. 5.12 *Left*: A twisted liquid crystal pixel transmits light. *Center*: When the voltage is turned on, the orientation of the liquid crystal changes and the light beam is blocked. *Right*: Different modes of orientation distortion: splay, twist, and bend (from *top to bottom*)

important part, not only in condensed matter physics, but in most fancy cosmological and quantum field theories. For instance, it may turn out that elementary particles are topological defects in the fine structure of the vacuum. In liquid crystals, defects obey the laws of topology as well, but they are tangible, and they can been seen and controlled.

Let us look first at textures in a thin nematic film where the director is oriented in the plane. The topological *charge* of a defect is measured by the rotation of the director along a surrounding contour. Stable defects with the lowest energy have the lowest possible charge. Since the director is invariant under rotation through 180°, the lowest charge is 1/2, either positive or negative. If a perpendicular orientation of the director is forced at the boundary of the film, the texture will appear as in the left panel of Fig. 5.13, where the lines show the local orientation. The total circulation along the boundary is one full turn, and therefore two half-charged defects remain at equilibrium. The nematic field near the defect is shown in the center of the same figure. Negative half-charged defects have a different three-fold structure. They appear when holes are cut in the film, as in the right-hand panel of Fig. 5.13, shaded in the way it would be seen through a polarizer.

In a bulk three-dimensional nematic, one can see line defects with a local structure in their cross-section similar to what was just described. There are also point defects, called *hedgehogs*, suggestive of their shape, as sketched in Fig. 5.14. A tale

Fig. 5.13 *Left*: A texture in a nematic film with two positive half-charged defects. *Center*: The director field near positive (*top*) and negative (*bottom*) half-charged defects. *Right*: The nematic field with two negative half-charged defects in a film with two holes

of hedgehogs, which I tell here without referring to names, illustrates the difference between the approaches used by mathematicians and physicists. To a certain approximation, when splay, twist, and bend elasticities are assumed equal (which is not likely to be true, but otherwise nothing can be done analytically), the energy of a hedgehog can be computed exactly. Moreover, the energy does not change if the hedgehog is "combed", as shown in the central panels of the figure. If we comb the hedgehog to the very end, so that all "needles" go in the same direction, we can connect the defects in pairs, as in the right-hand panel of Fig. 5.14, taking care that the total length of these lines is minimal. All hedgehogs remain in place, and the total energy concentrated in the connecting lines is minimized.

There is an entire subfield, with dozens of papers citing one another, but largely disconnected from the outside world, called "heat flow on harmonic maps", which studies textures in this way. Why heat? Because the equations used are similar to those describing temperature distributions. Why harmonic? Because the solution is obtained using harmonic functions of a complex variable. Why maps? Because mathematicians are astute in mapping things one upon the other, just as the surface of a sphere is mapped onto a plane sheet, but in more diverse and sophisticated ways. Why is the above conclusion utterly wrong? Defects have an internal structure. When the change of orientation becomes too sharp, nematic order disappears.

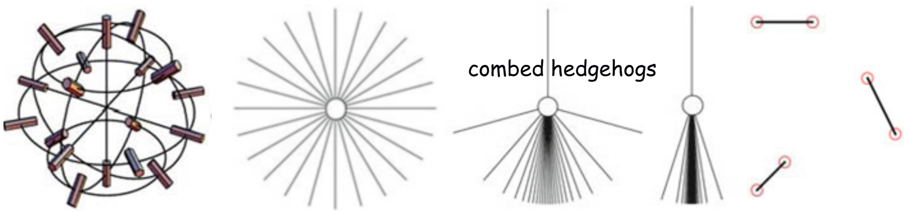


Fig. 5.14 *Left*: A three-dimensional rendering of the nematic field around a hedgehog defect. *Center panels*: Combed hedgehogs. *Right*: Mathematician's version of equilibrium, retaining six paired hedgehogs combed to perfection

This happens within the defect core, on a nanometer scale called a healing length. The core has an energy of its own. The physicist takes it into account and tries to match the core to the outside texture. This causes defects to move. Hedgehogs do not have a definite sign and are all attracted to each other. The connecting lines in the right-hand panel of Fig. 5.14 carry energy and tend to shrink, so that defects annihilate. Annihilation can still be prevented if the director must be oriented in a certain way on confining walls, but this is another story.

This narrative is just a brief glimpse. Entire books can, and have been, written on defects in liquid crystals, even on their equilibrium (Kleman, 1983), and their dynamics is still poorly understood.

5.6 Polymers

Some molecules can be huge: these are polymers. The story of synthetic polymers began in the mid-19th century, and Jacobus Henricus Vant Hoff, the first laureate of the Nobel Prize in Chemistry, is widely referred to as the grandfather of polymer science (Patterson, 2012). Polymers are synthesized by joining together monomer molecules, most commonly identical (Fig. 5.15 upper). In the simple version, polymer solutions are described in a way similar to the van der Waals theory, including just entropic terms and polymer–solvent interactions (Flory, 1953). Flexible polymer chains bend randomly, and a polymer molecule takes the form of a loose spherical blob (Fig. 5.15 lower left). At high concentrations, polymer chains become entangled (Fig. 5.15 lower center). Such an entangled state would already be a soft

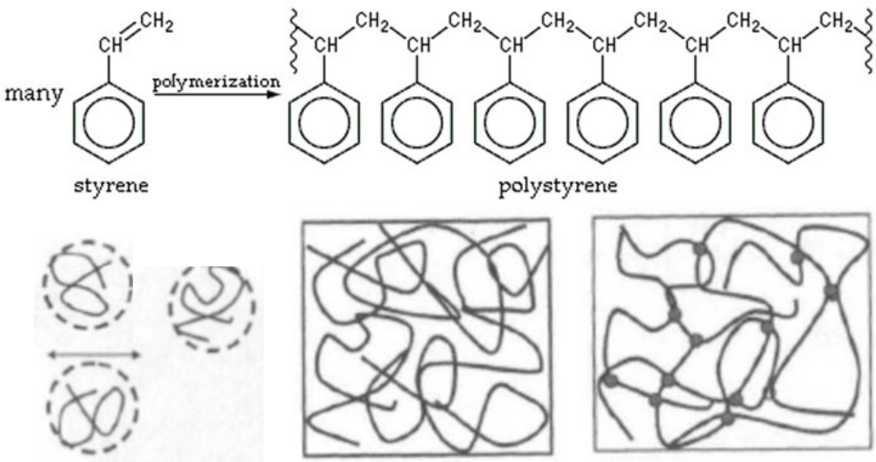


Fig. 5.15 *Upper*: A polymerization scheme. *Lower*: Polymers in dilute (*left*) and concentrated (*center*) solutions, and a cross-linked gel or rubber (*right*)

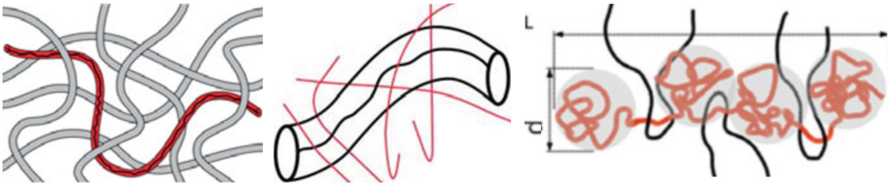


Fig. 5.16 *Left*: A reptation scheme. *Center*: A tube for a reptating chain. *Right*: Blobs between entanglements

solid if the solvent were removed, but to harden it, the chains should be cross-linked (Fig. 5.15 lower right). This is how rubber is made in the vulcanization process.

The dynamics of entangled polymer chains determining their mechanical properties is extremely complicated. Pierre-Gilles de Gennes was awarded the 1991 Nobel Prize in Physics for his ingenious theories of complex condensed matter – liquid crystals and polymers. His way of describing how entangled chains move is *reptation* (De Gennes, 1971), a suggestive term making us think of snakes. They creep as if in a tube between other chains (Fig. 5.16 left). Very long chains may form blobs, not unlike those in solutions when entanglements are relatively rare (Fig. 5.16 right).

Polymers and liquid crystals come together in a medium with unusual properties – liquid crystal elastomers, synthesized by polymerizing liquid crystalline molecules or attaching them to polymer chains. This was also an idea of de Gennes (1975), envisaged as artificial muscles. When this material undergoes a phase transition from the isotropic to the nematic state, it elongates significantly along the director and accordingly shrinks in the normal directions to preserve its volume; it goes in the opposite way when the transition is reversed. In this way, it really can work as a muscle lifting a considerable weight, but much more can be done by reshaping a flat nematic sheet in different ways (Fig. 5.17). By repeating phase transitions back and forth under the action of light or chemical agents, these soft shells can be made to walk and swim, and it may be possible to use liquid crystal elastomers to construct soft robots.

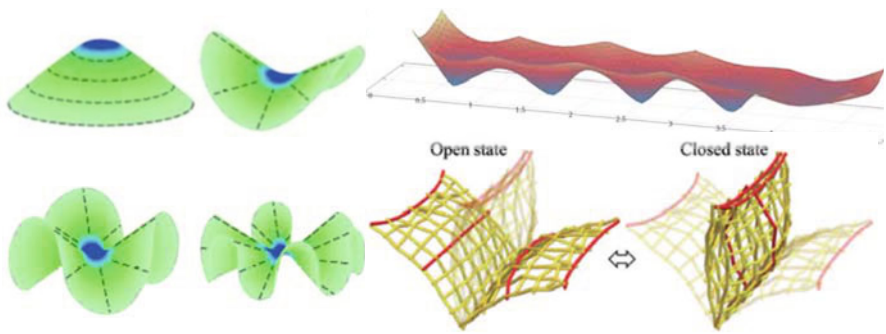


Fig. 5.17 *Left*: Forms obtained by deforming a flat disk. *Top right*: A walker with alternately flattening conical legs. *Bottom right*: A snapping trap made of a cloth with woven nemato-elastic fibers

Chapter 6
Quantum Matter

6.1 The Need for Quanta

The turn of the 20th century was a time of great change in social attitudes and arts, and it bore the seeds of the madness of the "short 20th century" 1914–1989, and in particular, of its first part 1914–1953. It was also a time of great progress in science. But it was only by a stroke of good luck that some consequences of these discoveries did not cut the entire story short.

Even though atomic theory was not yet universally accepted in the late 19th century, both physicists with their kinetic theory of gases and chemists with their models of molecular structure were regularly dealing in atoms. Atoms themselves were, however, indivisible and therefore inscrutable, since we need to take a thing apart to understand how it works. Nobody was audacious enough to try this, but the atom revealed itself by chance. Henri Becquerel (1896) was studying the phosphorescence of uranium salts in sun rays, hoping to detect the emission of X-rays recently discovered by Wilhelm Röntgen. The days were cloudy, and he left a wrapped photographic plate with an inserted cut-out metal screen next to a salt sample in a dark drawer. When he developed the plate, he was surprised to see vivid silhouettes of the metal screen. He recognized that radiation was coming from the uranium itself. This was the discovery of radioactivity, triggering further work by Pierre Curie and Marie Skłodowska-Curie, who isolated strongly radioactive polonium and radium (Curie et al, 1898). All three shared the 1903 Nobel Prize in Physics.

A matching piece was the discovery of the electron by J. J. Thomson (1897), which brought him the 1906 Nobel Prize in Physics. Starting from the 18th century, there were various speculations about whether positive and negative electric charges were due to the presence of two electric fluids. Faraday's concept of positive and negative ions carrying current in solutions (see Sect. 3.4) clarified this problem. Since ions had a certain valency, the charges could be naturally quantized, provided that the atomic hypothesis was accepted. George Johnstone Stoney (1881) came up with the idea of an "atom of electricity", which he called the *electron*, bearing a unit charge equal to the charge of a univalent ion. However, he assumed it to

L. Pismen, *The Swings of Science*,
https://doi.org/10.1007/978-3-319-99777-3_6

be permanently attached to atoms. The actual discovery, as well as the discovery of X-rays, came in the aftermath of the 1870s fashion of studying electrical conductivity, spectra, and emissions in tubes containing rarefied gases. J. J. Thomson identified "cathode rays" observed in these experiments as particles rather than waves, and measured their mass-to-charge ratio, which turned out to be independent of the cathode material. Becquerel (1900) found that beta-rays emitted by radioactive materials were deflected by an electric field and had the same mass-to-charge ratio as J. J. Thomson's particles. The term *electron* was adopted and kept for good, and indeed the suffix *on* was later attached to the names of all the elementary particles.

J. J. Thomson came next with the "plum pudding" model of the atom, where electrons are distributed like "plums" within a positively charged continuum (Fig. 6.1). From the point of view of classical electrodynamic theory, this actually made more sense than the later "solar system" model: if the distribution of electrons were disturbed it would be restored by forces of attraction to the positively charged bulk of the atom. Ernest Rutherford, his protégé, set about confirming this model by bombarding a gold foil with positively charged alpha-rays[1]. He was in for a surprise. Most of the alpha-particles passed straight through the foil, but a few were deflected through a large angle, which would have been impossible if the "plum pudding" model had been true. Aristotle would have been appalled: matter turned out to be almost totally empty, with atoms containing just a tiny nucleus occupying about a millionth of a billionth (10^{-15}) of its volume[2], surrounded by a coterie of still tinier electrons in the void.

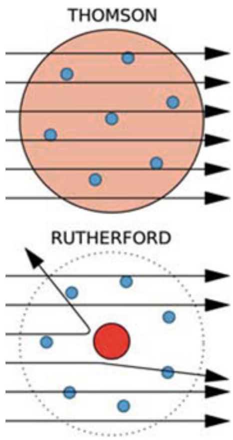


Fig. 6.1 J. J. Thomson's "plum pudding" model (*top*) and Rutherford's nuclear model (*bottom*)

Rutherford's "solar system" model inspired profane imagination. A superficial but fashionable Russian poet wrote: "perhaps the electrons are worlds similar to the Earth". Just imagine! They would have had their own electrons, *ad infinitum*, like Jonathan Swift's sequence of fleas feeding on bigger fleas. Lenin, in a philosophical mood, observed that the electron was as inexhaustible as the atom. But electrons are far from being inexhaustible, they are simple, and each electron is like every other one. It would be easier to build communism if people were like this. In a way, superconductivity (Sect. 6.6) is a kind of communism for electrons; one only needs to cool down their community to suppress thermal randomness. Atoms are a bit more complicated but not much, and superfluidity is like communism for atoms. Reality is more interesting, as Nature does not copy itself on different scales, as we already noted in Chap. 1.

[1] As the nature of radioactive emissions was originally unknown, they were given arbitrary names, viz., alpha, beta, and gamma rays, later identified, respectfully, as helium nuclei, electrons, and short-wave electromagnetic radiation.

[2] "Like a gnat in the Albert Hall", in Rutherford's words.

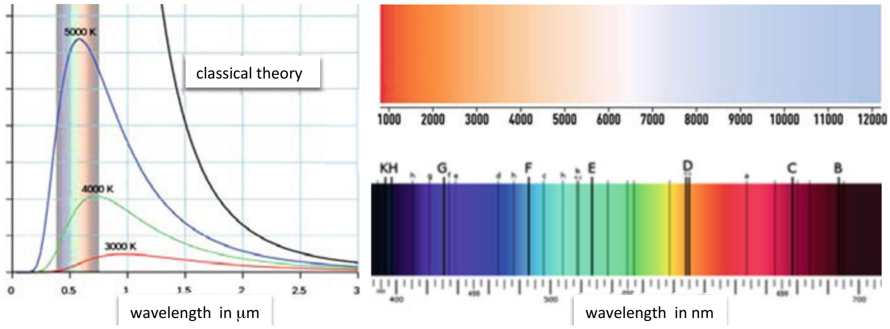


Fig. 6.2 *Left*: The intensity of black body radiation as a function of wavelength according to Planck's formula. *Top right*: The color of black body radiation as a function of temperature (in degrees kelvin). *Bottom right*: Fraunhofer lines

After the quantum of charge came the quantum of energy. It started quite innocuously from a strange paradox. According to classical thermodynamics, the energy of a black body should be equally distributed among all its degrees of freedom, which can be represented as Fourier modes of all frequencies and corresponding wavelengths. This is because a black body absorbs all radiation like a black hole but, unlike a black hole, it also emits radiation and so can (theoretically) attain equilibrium with its surroundings. This amounts to its having an infinite heat capacity and the radiated energy diverging according to the Rayleigh–Jeans law, in proportion to the square of the frequency, whence all such bodies should glow blue and emit copious amounts of the then recently discovered X-rays and gamma-rays.

Max Planck (1900) corrected this obvious nonsense by finding a way to suppress high frequencies; his formula included what came to be called the Planck constant, arguably the most important fundamental constant in the Universe. Planck's formula gave a reasonable result with the emission reaching a maximum at a certain frequency that increases with temperature (and hence at a wavelength that decreases with temperature) and decaying to zero at both extremes (Fig. 6.2 left). This conforms perfectly to common sense: we know that as the temperature of a metal rises, it first radiates heat, i.e., infrared waves, then glows red, then yellow, etc. (Fig. 6.2 top right). However, the Planck constant has the dimensions, not of energy, but of *action* (energy multiplied by time). Hence, to obtain the quantum of energy, cited in the justification for his 1918 Nobel Prize, we have to multiply it by the radiation frequency. The classical (Rayleigh–Jeans) formula is recovered when the ratio of this quantum of energy to thermal energy goes to zero.

The same energy quantization became the basis of Einstein's theory of the photoelectric effect (Einstein, 1905a). Light can knock out electrons when shone on a material. However, this can only happen when the wavelength of the light is shorter than a certain value, dependent on the material. The shorter wavelength, the higher the energy of the emitted electrons, while the light intensity affects only the intensity of emission. Einstein explained this by assuming light to be a stream of particles –

photons – with energy proportional to their frequency, each carrying Planck's quantum of energy. This was in some sense a revival of Newton's corpuscular theory of light – but it did not deny the prevailing wave theory, since the energy of a photon was related to the length and frequency of the wave. Light was now understood as both a wave and a particle.

Another hint at quantization, quite evident but not understood for a long time, came from observations of discrete spectral lines. William Hyde Wollaston (1802) noted dark gaps in the solar spectrum, later independently discovered and systematically studied by Joseph Fraunhofer (1814) and subsequently named after him (Fig. 6.2 bottom right). Gustav Kirchhoff (1859) identified some Fraunhofer lines with spectral emission lines of certain elements and deduced that the dark lines were caused by absorption by the same elements in the solar atmosphere. Johann Balmer (1885) derived a simple empirical formula for the lightest hydrogen atom. The main message, still undeciphered, was the *discreteness* of these spectra.

Niels Bohr (1913) united the ideas of Planck and Rutherford with the discreteness of atomic spectra in his audacious model of the atom. It envisaged the electrons moving around the nucleus along a fixed discrete set of orbits with the radii related to one another as squares of natural numbers. The differences between the energies of these orbits precisely corresponded to the observed spectral lines of hydrogen atoms. However, this picture defied the laws of classical electrodynamics, which predict that an electron with this motion should continuously lose energy, and eventually fall onto the nucleus. This apparent defect was retroactively vindicated by quantum mechanics, as we shall see presently.

6.2 Quantum Weirdness

It was a short step from Einstein's viewing the photon as both a wave and a particle to extending this duality to electrons, as conjectured by Louis de Broglie (1925). The wavelength of the associated wave turned out to be about the same size as the hydrogen atom. De Broglie came up during his long life with several theories as scantily justified but wrong, and hardly deserved being awarded the 1929 Nobel Prize before the actual creators of the "new" quantum theory.

Proper quantum mechanics was formulated in two equivalent ways that seemed incompatible at first sight, by Werner Heisenberg (1925) and Erwin Schrödinger (1926). The Schrödinger equation, revolutionary as it has been, is quite simple. It resembles the heat equation, but the time derivative comes with the imaginary unit, which makes it reversible, and the interaction energy stands in the place of a heat source. This algebraic term is actually the most complicated one, as it involves the interactions with all the relevant particles and fields. What is most remarkable is the dependent variable, the wave function, replacing temperature. It is a complex function that has no physical meaning by itself, but its squared absolute value (modulus) defines the probability of finding the particle at a given location at a given moment of time.

Heisenberg worked in a more refined and unconventional way. He defined positions and momenta of particles not just as numbers as in classical mechanics but as *operators* acting on the parameters defining the state of a particle. Operators, unlike numbers, generally do not commute. Thus, if we transpose the coordinate operator, which is just the coordinate, and the momentum operator, which is the derivative with respect to this coordinate, the result will obviously change. This led to Heisenberg's uncertainty principle: any two variables can be measured simultaneously only when the respective operators commute. The Schrödinger equation can be built up out of the above operators but Heisenberg represented his operators in another way, by non-commuting matrices, re-inventing matrix algebra along the way, although he was not aware of this; originally he was even confused by the failure of his operators to commute. It was therefore not evident from the outset that the two approaches were equivalent. Heisenberg and Schrödinger clashed, and Schrödinger was appalled by the probability interpretation of his wave function when it was first suggested by Max Born. The great men were just unable to grasp what is now taught to undergraduates – till everything finally got sorted out.

Bohr was so enamoured of Heisenberg's uncertainty principle that he extended it to the "complementarity principle", supposedly applicable to all phenomena, both natural and cultural. Other founders of quantum mechanics followed suit. Max Born (1962) asked: "There must exist a relationship between the latitudes of freedoms and of regulation [...] but what is the political [Planck] constant?" Connections between quantum mechanics and psychic phenomena were broached. This was gleefully embraced by postmodern nonsense philosophers who were later parodied in a famous hoax by Alan Sokal (1996).

A constellation of young theorists congregated around Niels Bohr's Institute in Copenhagen. Besides Heisenberg, these included Wolfgang Pauli, Paul Dirac, Lev Landau, and George Gamow. The more mature participants of this drive called it "Knabenphysik" (boy's physics). Quantum mechanics worked like a charm, cracking every problem in sight, but the question was (and still is): what does it mean? Bohr said, "there is no quantum reality", Wigner said, "there is no reality without an observer", and Heisenberg said, "the idea of an objective real world whose smallest parts exist objectively in the same sense as stones or trees exist, independently of whether or not we observe them is impossible".

The state of a quantum system is unknown until the wave function "collapses" upon observation. Any observation, indeed, disturbs the state of the system. For example, if we wish to detect the particle's position with higher and higher precision, we need to use light with a shorter and shorter wavelength – but then the velocity of the particle absorbing the energy of the detector's ray becomes less and less definite. As the joke goes, a traffic policeman asks Heisenberg: "do you know how fast you are driving?" He answers: "no, but I know where I am". But does this mean that the particle does not have any properties before they are measured? The extreme view is the "participatory Universe". Perhaps the world would not exist if we were not here? But isn't that the good old solipsism, a convincing if extremal justification of the anthropic principle? In a splendid 1964 poem, Joseph Brodsky (the 1987 Nobel Prize in Literature – not Physics) sent a long farewell, while his boat was sinking in

Fig. 6.3 Schrödinger's cat

the Finnish Bay, to remote lands he never visited and to Newton, Freud, Einstein, and others whom he only dreamt of, and to the entire world: he knows that all this ceases to exist as soon as "the Old Lady switches off the light". John Wheeler (1998) thought of the Universe as a self-exciting circuit, existing because we exist, but apparently, unlike in Brodsky's poem, due to a collective effect, still continuing after Wheeler himself passed away.

Wheeler explains the Copenhagen interpretation, the standard doctrine promulgated by Bohr, in his informal way (somewhat rephrased here). You drive down a road and come to a fork. According to classical physics, you take one of the two roads, and that's that. According to Bohr, you may take one fork or the other, and it will be unknown until you stop at a gas station and an outside observer ascertains your location. "There is something ghostly about it", writes Wheeler, "as it assumes that you travel virtually down both roads at once" until apprehended. It is, indeed, most weird when applied to macroscopic objects, like this driver. The same kind of an example, but more cruel, was Schrödinger's cat (Fig. 6.3). A cat is enclosed in a steel box, and a radioactive signal, which may or may not be produced by a single atom decaying, triggers the release of a poison. Until the box is opened, the cat remains in a superposition of alive and dead states. Schrödinger did not dare to subject a human to this thought experiment.[3] But after all, a cat is a conscious being too. Could it not serve as an observer itself? And Wheeler's motorist, too? Sheldon Goldstein (1998) writes: "Many physicists pay lip service [. . .] to the notion that quantum mechanics is about observation [. . .] but hardly anybody truly believes in this anymore – and it is hard for me to believe that anyone really ever did".

[3] There was a bizarre proposal to abolish inhuman incarceration and instead to place a convicted felon on the electric chair for a certain time, with the deadly jolt being administered at random with a probability matching the gravity or the crime.

6.3 Objections and Bypasses

Einstein deplored the probabilistic overtones of quantum mechanics saying that "God doesn't play dice", and tried to invent counter-examples proving its inconsistency. The most successful attempt was the Einstein–Podolsky–Rosen (EPR) paradox (Einstein, Podolsky, and Rosen, 1935). There are two entangled particles, electrons or photons. Their entanglement means that some property of each particle, say, polarization, is unknown, but it is known from the way these particles have been created that, say, one of them should have left polarization and the other one right polarization. The particles go their separate ways carrying their superposition of left and right states, but when one of them is caught and witnessed to have right polarization, the other, faraway one, suddenly becomes definitely left-polarized. Intended as an undoing of quantum mechanics, the EPR effect turned out to be real, and is being proven to this day by experiments at longer and longer separations between entangled particles. The sting of the EPR paper was in the action at a distance being incompatible with relativistic causality, which forbids communication with any speed faster than the speed of light. However, it has now been established that it would be impossible to use the EPR effect to transmit information.

Einstein was not alone to suspect that there should be something hidden beyond the probabilistic veil of quantum mechanics. Erwin Madelung (1926) came up with a hydrodynamic formulation of the Schrödinger equation resembling Euler's equations for a compressible fluid, with the density representing the probability distribution, the velocity related to the probability current, and the pressure replaced by the chemical potential taking account also of all interactions. There was no mistake in all this, but it did not eliminate the probabilistic description and brought no computational advantages. The main direction was finding some hidden variables. John Bell (1964) devised tests proving that *local* hidden variables were incompatible with experimental observations. This left a possibility of non-local hidden variables, the most advanced attempt of this kind being the "pilot wave" theory suggested and abandoned by de Broglie and later developed by David Bohm (1952).

Instead of the particle–wave duality of mainstream quantum mechanics, de Broglie and Bohm envisaged *both* a particle moving along a classical trajectory and a pilot wave that would interact with it. By the standard interpretation, a delocalized particle passes through both interstices in the double-slit experiment, and the interference picture appears because its wave function has two sources at the two slits. By Bohm's theory, the particle passes through a certain slit, but its interaction with the pilot wave causes the interference pattern. This alternative was realized on the macroscopic level by Yves Couder and Emmanuel Fort (2006). It started from the experiments of Couder's group showing that a millimeter-sized droplet does not coalesce when falling on a vertically oscillating layer of the same liquid but jumps off and turns into a "walker" guided by the Faraday waves on the liquid surface (Sect. 7.1). The interaction with the waves, which depends on the phase of the wave at the moment of collision, gives the walkers some kind of non-locality reminiscent of the quantum-mechanical picture. When they are directed through a double slit, it would be impossible to deduce from the interference pattern which slit the walker has passed through if it had not actually been seen. The differences between this macroscopic pilot wave experiment and its quantum-mechanical analog are, however, too great to give support to the microscopic pilot wave theory.

Parallel to the honest attempts at understanding and improvement, the baffling ideas of quantum mechanics and relativity were attacked both from the left and from the right, and from below, as they did not mix well with the totalitarian ideologies gaining strength in the 1930s or with common boorishness. They were decried as "Jewish physics" by the Nazis and as "idealistic concoctions" by Stalin's ideologists. The Nobel Prize laureates Philipp Lenard and Johannes Stark were about the only prominent scientists who promulgated the non-existent "Deutsche Physik". We know what this meant: Germany lost its leadership in physics to the USA. Heisenberg kept quiet and either sabotaged or was unable to advance the German nuclear bomb program which he headed (another Heisenberg uncertainty, perhaps), while illustrious emigrants eagerly took part in the Manhattan project. Stalin was more astute. When adepts of materialistic science came up with a list of idealistic physicists, he showed it to Igor Kurchatov, the head of the nuclear project[4]. Without them, Kurchatov said, there will be no bomb. This was the end of the story.

The most radical attempt to avoid quantum uncertainties was the "many worlds" hypothesis of Hugh Everett III, presented in his 1957 PhD dissertation. Going back to Wheeler's simile (Wheeler, 1998), according to the "many worlds" interpretation, after you stop and meet an observer you are yourself aware of being there, but this does not mean that there isn't another decoupled "you" that stops to eat at another road junction and meets other people and becomes aware of being there. Some reasonable people seem to believe this [see, e.g., Tegmark (2014)], but there are just too many forks that can be encountered in a continuous manner, and not only by yourself, but by other people and other creatures and all kinds of quantum particles, even in remote galaxies: and they may all affect you. There are just too many universes! No one could ever count them, and perhaps they could be assigned Cantor's number Aleph-Infinity. Neither the reader nor myself can imagine what this is. The continuum is Aleph-1, and nothing higher is known to exist. Georg Cantor believed his theory had been communicated to him by God – who may dwell in Everett's multiverse. Wheeler, who was Everett's PhD advisor, did not accept his crazy ideas but found them interesting and not only helped Everett to get his PhD, but presented him to the great patriarch Bohr himself. Bohr was, of course, appalled. Everett, whose life's work outside science after this PhD was evidence of his exceptional talents, believed in quantum immortality among the uncountable multitude of worlds, and perhaps because of this was drinking heavily and chain smoking, neglecting the health of the accidental body he happened to occupy on this insignificant stretch of road, and died on this Earth in his early fifties.

The interest in the "spooky" aspects of quantum mechanics and its spurious spiritual connections was renewed with the blooming of "counterculture" in the 1960s and 1970s. Kaiser (2011), describing the "Fundamental Fysiks Group" formed by eccentric outsiders to "save physics", contends that they "planted the seeds that

[4] His huge disembodied head is now standing at the gate of the institute bearing his name, while countless statues of Stalin have been broken up or melted down.

would eventually flower into today's field of quantum information science". But physics needs no saviours; it is not doing so bad, even though it lacks breakthroughs on the scale of those occurring in the early 20th century, and "quantum information" is not its most glorious field. The most fruitful attitude to quantum weirdness has been "shut up and compute". Whatever way the wave function is interpreted, it evolves deterministically and faithfully describes all observable phenomena. Perhaps some questions just need not be asked because they are incompatible with Nature's language.

6.4 Chemistry Explained

The easiest but most momentous computation was solving the Schrödinger equation for the hydrogen atom. Reportedly, Schrödinger could not do it himself but was helped by a friend mathematician. This was an exercise in separation of variables in spherical coordinates, leading to solutions with the angular dependence expressed through spherical harmonics. The solutions are tagged by three natural numbers: the first, n, entering only the radial function, the second, ℓ, affecting both radial and angular dependences, taking integer values between 0 and $n-1$, and the third, m, affecting only the angular function, taking integer values between $-\ell$ and ℓ. The shapes of some spherical harmonics, labelled in a standard way we will talk about presently, are shown in the upper left panel of Fig. 6.4. This is just a sketch, showing the symmetry but not the radial distribution of the probability cloud. Chemists wondered why the first row of Mendeleev's table has two elements, the following two, eight, and the next, eighteen (at least, as it is commonly displayed now). The sequence looks suspiciously like squares of 1, 2, 3 doubled. The structure of this family of solutions, with proper corrections, explains why.

Before proceeding further, we need to recall the existence of *spin*, the internal angular momentum of a particle. Electrons, as well as protons and neutrons, have spin 1/2. All electrons, as well as other quantum-mechanical particles of the same kind, are indistinguishable, unlike Couder's droplets from the preceding section, which the researcher could have colored blue and red if he had so wished. Particles with half-integer spin are called *fermions*. They obey Fermi statistics: if you interchange two particles, the wave function should change its sign. This would be impossible if these particles were in exactly the same state. Therefore fermions obey the Pauli exclusion principle: each quantum-mechanical state should be occupied by a single particle. Photons, unlike matter particles, have spin 1 and obey Bose–Einstein statistics: when they are interchanged, the wave function remains the same. This allows us to enjoy intense colors given by plenty of identical photons with the same frequency.

Electron orbits can be vividly represented by the atomic theater of Fig. 6.4. The solutions of the Schrödinger equation have different energies growing with increasing values of n. The first electron entering the theater takes the best seat in the first row, $n = 1$. This is the hydrogen atom, #1 in the Periodic Table. There is only one

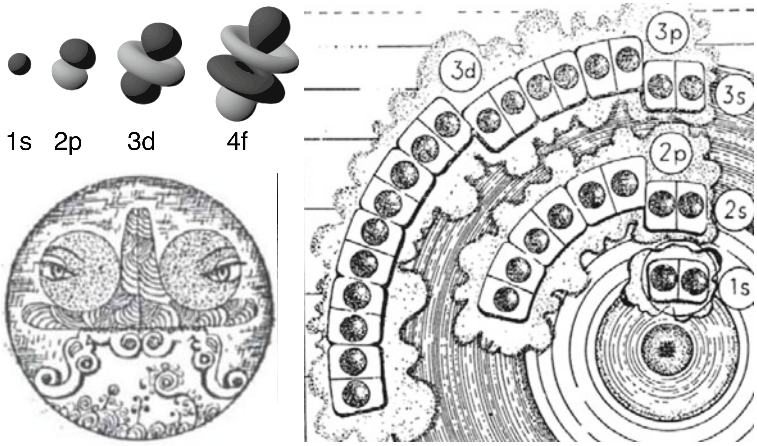


Fig. 6.4 *Top left*: Examples of shapes of electron clouds. *Bottom left*: A double bench for electrons with antiparallel spins. *Right*: The atomic theater

bench available in this row, with $\ell = 0$. The corresponding orbit is spherically symmetric and is labelled by the letter s. The bench accommodates two electrons with antiparallel spins (see how they sit in the lower left panel of Fig. 6.4). The second electron of the helium atom, #2, grabs the remaining seat. The row, or in the standard terminology, the electronic shell, is closed. Atoms with closed electronic shells are satisfied with their stable state and reluctant to enter chemical reactions; helium is the first of the noble gases.

The next row already has four benches. The best one is symmetric, $n = 2$, $\ell = 0$, labelled $2s$. Three others, with $\ell = 1$, are labelled $2p$. The value of m affects only the orientation of the electron cloud, but not its energy. Eight more electrons take the seats in the second row, bringing us to #10 in the Periodic Table, another noble gas neon. In the third row, with $n = 3$, there are again four $3s$ and $3p$ benches filled up to #18 argon. In addition, there are now five benches with $\ell = 2$ labelled $3d$. The electron of the next element, #19 potassium, grabs instead the symmetric $4s$ seat, which turns out to have a lower energy. This is a correction due to interactions between electrons, which are not taken into account by the simple solution. The $3d$ benches are filled up starting with #21 scandium up to #30 zinc, after which the $4p$ shell starts to fill up, up to another noble gas, #36 krypton; #37 rubidium prefers $5s$ to $4d$, and so it goes, with the seven $4f$ benches ($n = 4$, $\ell = 3$) only filled up by the lanthanide group, starting with #58: the further we go, the stronger the interaction effects among electronic clouds. Amazingly, Bohr's atomic model, even without the benefit of the Schrödinger equation, but solely based on intuition, predicted the structure of Mendeleev's table in the same way. Both Mendeleev and Bohr were daring visionaries.

Paul Dirac (1929) claimed confidently that "the underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known", so that he could now proceed with clear conscience to higher realms. The claim was repeated on a more solid basis by Linus Pauling who said in a 1992 interview to John Horgan (1996): "I felt that by the end of 1930s, or even in the middle, that organic chemistry was pretty well taken care of, and inorganic chemistry and mineralogy – except the sulfide minerals."

Quantum mechanics did indeed reveal the basis of chemical structure. Following the picture of Bohr's atom (still unsupported by the Schrödinger equation), the physical chemist Gilbert Lewis (1916) developed the concept of the electron-pair chemical bond. When the mathematical machinery of quantum mechanics finally arrived, it was still of little help. We have seen that even heavy atoms could not be handled quantitatively, let alone complex chemical molecules, but at least the nature of the dashes connecting the symbols of elements in the formulas of the 19th century chemists (see Sect. 4.4) were now understood. In a rough approximation, Heitler and London (1927) solved the Schrödinger equation for the hydrogen molecule – the simplest of all – and saw that two electrons with antiparallel spins come together between the two nuclei and tie them one to the other, as if a bench from Fig. 6.4 had been placed between two scenes. This confirmed Lewis's idea and solved the 19th century mystery of the attraction of similar entities.

The Schrödinger equation is an unwieldy key (Fig. 6.5) opening all the gates of chemistry, but turning it in the lock requires titanic strength. A picklock is more efficient, and this was Pauling's strategy. He described his approach to structural chemistry to Judith Goodstein (1984): "I try to identify myself with the atoms. [...] I ask what I would do if I were a carbon atom or a sodium atom under these circumstances." On a semi-quantitative level, it worked perfectly. The carbon atom has four electrons in the outer shell and Pauling (1931) felt that if he were a carbon atom he would like to "hybridize" the 2*s* and 2*d* electron clouds to point them symmetrically toward the vertices of a tetrahedron at an angle of 109°28' one to the other, and join suitably paired electrons of other atoms to each of them, whether it be carbon, hydrogen, nitrogen, or whatever is needed to build up variegated forms of organic molecules. This was indeed close to the actual angles between carbon bonds known from X-ray data. John Slater (1931) came up with similar qualitative ideas about directed bonds formed by *p*-electrons.

Fig. 6.5 The key to chemistry and picklocks

In the same year, Pauling cracked the structure of a double bond formed by two electron pairs, as in ethylene or molecular oxygen. A challenging problem was the structure of benzene, which could be assigned alternative structural formulas (Fig. 4.4). The breakthrough idea (Pauling and Wheland, 1933) was the *resonance* between the alternative Kekulé structures resulting in a reduction of energy relative to each of them separately. Pauling was awarded the 1954 Nobel Prize in Chemistry "for his research into the nature of the chemical bond and its application to the elucidation of the structure of complex substances". This was, of course, not the end of the story. More complex chemical bond structures exist, for example, in metal–organic complexes, and chemistry was not in the end brought to a close with Dirac or even with Pauling.

At this point, I would like to digress over another attack on science, back in the USSR, where an antiscientific and anticultural offensive was rampant in the dark postwar years, at a time when the Cold War almost heated up. Trofim Lysenko, who advocated a kind of home-grown Lamarckism and caused much damage to Soviet agriculture, but was much loved by Stalin, cracked down on "Mendelist–Morganists", proper geneticists who believed Gregor Mendel (more on this in Sect. 9.3) and, following Thomas Hunt Morgan[5], experimented with fruit flies instead of cows. At this time and place, scientific discussions were a very serious business: if you held wrong views, you were incarcerated or shot, as was the leading Soviet geneticist Nikolai Vavilov.

Inserting ideology into science got more difficult the further down the Marxist ladder of "forms of motion" one tried to do it. The highest was the social form, and here nothing could be said beyond the party line. The biological form was the next highest and also irreducible to lower forms, so it was almost as easy to reduce or raise to the level of Marxist ideology. Chemistry was already too far down in the darkness, a harder science and more resistant to ideological assault, but there were obscure chemists envious of Lysenko's success. Their easy target was Pauling's resonance theory. How come a molecule doesn't have a definite structure? What about materialism and objective reality? For these materialists, "reality" was contained in the dashes of familiar structural formulas rather than in electronic clouds. I mentioned above that quantum mechanics was objectionable as well, but it was even harder again to assail physicists.

A massive discussion meeting was convened in 1950. The putative chemical Lysenkos aimed high, in fact at the top chemists, right up to Alexander Nesmeyanov, the President of the Academy of Sciences, but they were out-maneuvered. In the end, everything was boiled down to just two "Paulingist–Ingoldist"[6] scapegoats: Yakov Syrkin, the corresponding member of the Academy and Stalin Prize laureate, and his collaborator Mira Dyatkina, both Jewish[7]. Even these traitors of materialistic chemistry were not jailed, and Syrkin even retained the chair of Physical Chemistry at the Institute of Fine Chemical Technology in a brick building that hosted the Russian analog of Radcliffe College before the revolution.

This institute was where I entered as an undergraduate in the footsteps of my father, whom I begrudged till the end of his long life the fact that he had not directed me to study physics instead. He knew that the complaints weren't serious, I was then more into poetry but thought that studying humanities in the Soviet Union was an oxymoron. I was mistaken, by the way; as the times were warming up, the mois-

---

[5] The winner of the 1933 Nobel Prize in Physiology or Medicine.

[6] In honor of Sir Christopher Kelk Ingold FRS, who in the 1920s and 1930s carried out important work on the electronic structure of organic compounds.

[7] When Jews were denounced at that time, they were not called Jews but "cosmopolitans"; everybody understood what that meant.

ture of high culture hidden in the pores of the regime started to seep out. I read the translation of Einstein and Infeld (1938), instead of Lenin's philosophical masterpiece, to prepare for freshman Marxism seminars; the lecturer loved my reports and passed me without an exam. In our junior year, Syrkin, by this time a full member of the Academy, launched an extracurricular seminar. Volunteers were given papers or book chapters on statistical physics and quantum mechanics to study and report; this was the best school.

In the early 1970s, I came across a bulky volume of proceedings of the 1950 discussion and published its story. By this time, Pauling was a beloved figure in the USSR, a long way from an "ideological warmonger", as he was called in the 1950s. His antinuclear activism (following work on explosives in war time) brought him both the 1962 Nobel Peace Prize and the 1972 Lenin Peace Prize, a sign of appreciation for his alleged communist sympathies, widely criticized in the US. All the same, the commissar chief editor, re-reading the journal issue, said she wondered how she had passed the article. The title was "In Memoriam for a Theory". The resonance theory that had so much disturbed the "materialists" was superseded by molecular orbitals. The electron cloud spreads out all over the benzene ring, shared by all six carbon atoms. In Israel I met a postdoc, a scion of the Russians who fled the revolution to China and further east to California, who translated my article for Pauling at his request.

After the early successes of these semi-quantitative theories, quantum chemistry turned into quite a boring discipline. At a certain stage, chemistry departments became the most voracious consumers of computer power on campuses, applying brute force to turn the heavy key of Fig. 6.5. Paraphrasing what Prandtl said about modeling flows (Sect. 2.3), the atom itself is the best, because it perfectly computes its force fields and spectra. Can it not be used for other computational tasks? Richard Feynman (1982) outlined the idea of a computer operating with single atoms based on quantum-mechanical laws. He is sometimes called the father of quantum computing, though less famous people expressed similar ideas at about the same time. This direction is now actively pursued. The quantum information unit is a *qubit*. Unlike a classical *bit*, which has the values 0 or 1, a qubit is in a quantum superposition of the 0 and 1 states, and much more extensive information is carried by its phase. Two qubits can be readily entangled, as in the EPR experiment (Sect. 6.3). At the 2018 meeting of the American Physical Society, Google presented the "Bristlecone" quantum computer with a record number of 72 qubits arranged like scales in a pinecone. Still, this is far from the complexity of a single atom, and miniaturizing working quantum computing devices is no easy matter. A major problem is sustaining qubits in a coherent state that retains their phase information. So far coherence is not long-lived, even at low temperatures. Some skeptics claim that quantum computers will never be reliable due to decoherence. Only time will tell.

6.5 Everyday Quantum Devices

A large part of the physics that Paul Dirac (1929) declared to be "completely known" is exactly what was subsequently developed most rapidly and brought most changes to our everyday lives in the late 20th and 21st centuries. Like chemistry, it is only concerned with the interactions of electrons in outer shells with atoms and radiation, and is actually simpler than the quantum chemistry of complex molecules.

Properties of both metals and semiconductors depend on the interactions of electrons with a lattice. Felix Bloch (1928) solved the Schrödinger equation for electrons moving in a periodic potential. The underlying mathematical theory of differential equations with periodic coefficients goes back to Gaston Floquet (1883). The spectrum of Bloch waves consists of bands occupied by electrons with a continuously changing wave vector. The continuous spectrum within each band is also quantized in a finite chunk of a material, and here comes Pauli's exclusion principle once again, allowing each quantum state to be occupied, like the seats in Fig. 6.4, by just one electron. The conductivity of a material depends on how easily electrons can move from one seat to another (Fig. 6.6 left). The boundary between energies of occupied and unoccupied states in the absence of excitations is called the *Fermi level*, shown by the dashed line in the picture.

If the highest occupied band, called the *valence* band, is fully occupied, while the energy gap separating it from the lowest unfilled band, called the *conduction* band, is large, the material is an insulator. Any insulator may, of course, suffer electrical breakdown when the applied voltage is high enough to set in motion some charge carriers, perhaps provided by ever present contaminants, or high enough to create them as when air turns into plasma in a lightning flash. At the other extreme are metals, where the valence band is not filled and electrons are easily excited to nearby levels and travel freely while becoming virtually shared by all the atoms in the lattice. Semiconductors stand in-between. They may have a few unoccupied levels

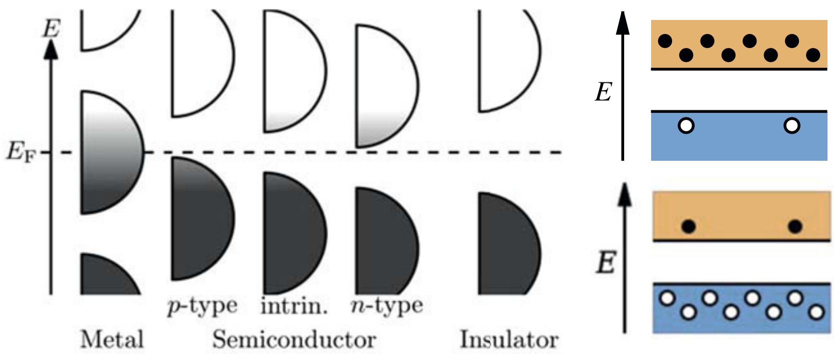


Fig. 6.6 *Left*: Change in the occupation pattern of electronic levels (dark when occupied) between metals and insulators. *Right*: Band structure of p-type (*upper*) and n-type (*lower*) semiconductors. *Black circles* in the conduction band are electrons and *white circles* in the valence band are holes

in the valence band; these are *p*-type semiconductors, where the remaining sites – positively charged holes – are the principal charge carriers. Alternatively, the charge carriers in *n*-type semiconductors are a few electrons in the conduction band. The in-between position is occupied by materials with a relatively small gap between the valence and conduction bands near the Fermi level; electrons can easily jump across this gap, becoming charge carriers themselves and leaving behind holes as additional charge carriers. The presence of charge carriers of two types is reminiscent of the 19th century discussion about whether there are two "electric fluids" or just one.

The main reason why semiconductors are so important is the ease with which their properties can be manipulated by doping. This is the way all components of modern electronics are manufactured. The material used to make the first working semiconductor devices, which were transistors to replace vacuum tubes in computers, was germanium; but soon industrial transistors climbed up a row in Mendeleev's table to silicon. Transistors coalesced into integrated circuits, manufactured in ever finer density and resolution by enormous lithographic machines in "clean rooms" of the size of large stadiums, where workers go around dressed almost like astronauts. "Silicon valleys" spreading from the San Francisco suburbs to many countries do not really deal with silicon, but with software feeding ever smaller and more powerful devices. Semiconductors graduated from the field of science to industry and information. They also moved directly to our homes, besides being packed into smartphones. In some semiconductors, excited electrons can relax by emitting light, something exploited in light-emitting diodes (LED), used to make bright low-energy-consuming lamps and television screens.

Another quantum device that has become a feature of everyday life is the *laser*. Its principle of operation is based on stimulated emission of coherent light. First of all, many electrons should be excited to a higher energy level to create a population inversion, defying the Boltzmann equilibrium distribution. The electrons are prevented from falling back spontaneously until stimulated externally, whereupon they all drop down, emitting photons with a phase, frequency, polarization, and direction of travel identical to the photons of the incident wave; the emission is enhanced by letting the resulting *coherent* light wave circulate within a medium between two mirrors (Fig. 6.7). It is the precision of the acutely monochromatic coherent wave that makes lasers so valuable for numerous applications.

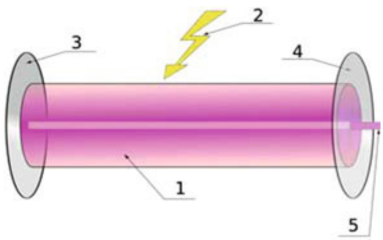


Fig. 6.7 Components of a typical laser: (1) gain medium, (2) energy pumping, (3, 4) reflectors, (5) laser beam

Who should really be credited with the invention of the laser and its long-wave precursor, the maser, is not clear, in spite of the Nobel Prizes handed out. The theoretical principle of stimulated emission goes back to Einstein (1916a). The Moscow professor Valentin Fabrikant started to work on its realization in the late 1930s and in 1951 filed a patent application, granted in 1959, followed by a discovery certificate in 1964, as retold in his obituary (Biberman et al, 1991). His work was far from being obscure (I even heard of it from my then mother-in-law, a lady excelling in academic gossip), and it could not have been unknown to Nikolay Basov and Alexander Prokhorov, the communist establishment figures who shared the 1964 Nobel Prize in Physics with Charles Townes, which did not prevent bitter clashes between them.

Arthur Schawlow, winner of the 1981 Nobel Prize in Physics for his contribution to the development of laser spectroscopy, whose joint work with Townes gave the theoretical basis for the first American laser (Schawlow and Townes, 1958), reportedly said once that lasers would never have any practical significance. Powerful rays bringing down planes and rockets, imagined by science fiction writers even before the invention of lasers, are still not operational at the time of writing, but lasers are all around us, in laser printers, laser surgery, fiber optic communication, and measuring all kind of things, from the distance between apartment walls to the distance to the Moon to the speed of your speeding car. Some uses are already becoming obsolete as laser disks give way to direct downloads of noisy pop music from the cloud. And even without the benefit of high power, houligans may have an opportunity to bring a plane down by blinding a pilot with a simple laser pointer.

6.6 Quantum Collective Effects

Quantum condensed matter is no less rich and complex than the outer reaches of physics capturing public attention, and often displays even more beautiful symmetries, singularities, and phase transitions. They manifest themselves in collective phenomena involving macroscopic collectives of quantum particles.

After being the first to liquefy helium (which brought him the 1913 Nobel Prize in Physics), Heike Kamerlingh Onnes (1911) discovered superconductivity in mercury. Nobody had ever experimented at such low temperatures before, and the conductivity was expected to decay to zero as all movement was arrested close to the absolute zero temperature. But rather than the conductivity, it was the resistance that vanished. It took almost half a century to understand this effect.

The next low-temperature discovery was superfluidity. Pyotr Kapitza was carrying out low-temperature studies in the Cavendish Laboratory, but was not allowed to go back there after a home visit to the USSR in 1934, and could only continue research after equipment was moved to Moscow. He observed the superfluidity of liquid helium in 1937 (Kapitza, 1938). Meanwhile, John Allen and Don Misener, who continued this investigation where Kapitza had been working in Cambridge, came upon the same discovery and published it in the same issue of Nature (Allen and Misener, 1938). Priority disputes delayed the Nobel Prize – apparently the two young physicists were not judged worth it, and Kapitza alone was awarded one for his life's work in 1978, sharing the cash, in the Swedish Academy's sophisticated ways, with the discoverers of the cosmic microwave background radiation.

Superfluidity was not as mysterious and unexpected as superconductivity. Helium atoms, with their pairs of protons, neutrons, and electrons, are bosons obeying Bose–Einstein statistics; unlike fermions, they do not ask for individual chairs but can all crowd into the same lowest energy state – the Bose–Einstein condensate (BEC). As of 1995, gaseous BECs were observed at nano-kelvin temperatures, and the field blossomed, rediscovering some phenomena previously observed in superfluid helium and finding more exotic patterns.

Superfluid helium is in some way the embodiment of the ideal fluid, the 18th century "dry water". Although superfluidity is a quantum phenomenon, it involves a macroscopic number of atoms so it can be described phenomenologically using the universal equations due to Landau (1937), which are applicable to all kinds of phase transitions (Sect. 5.2). The order parameter describing the superfluid condensate is a complex scalar, and the nonlinear equation derived by Landau's rule bears the names of Eugene (not the more famous David) Gross and Lev Pitaevskii, the coauthor of all the books in the famous Landau–Lifshitz series published after Landau's debilitating accident. The connection with hydrodynamics is not obvious, but it is made clear in the same way as in Madelung's hydrodynamic formulation of the Schrödinger equation (Sect. 6.3): the squared absolute value of the order parameter is identified with the fluid density and the gradient of its phase with the fluid velocity.

Vortices appear in a superfluid as in a classical ideal fluid. They look and behave in a similar way, but now they are upgraded to topological singularities of the order parameter field, and the flow circulation around them is quantized; within narrow vortex cores, the superfluid density goes down to zero. If a vessel containing the superfluid is rotated, vortices arrange themselves parallel to the rotation axis, as in Fig. 6.8. Each vortex carries one quantum of circulation, and the faster the vessel is rotated, the more vortices appear; the rotating vortices, repelling each other as electric charges do, arrange themselves in a regular lattice. Vortices will also appear easily in a superfluid agitated in any way, forming ever changing tangles. As usual, nothing is really ideal; there is always an admixture of normal fluid helium, and the Gross–Pitaevskii equation, as well as Euler's equation, are just nice models that should be corrected by dissipative effects if it is essential to know the precise flow pattern.

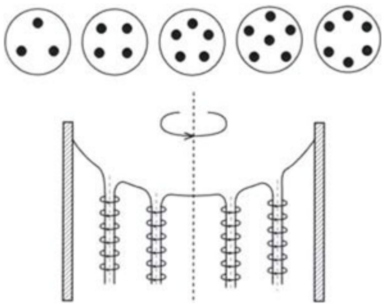


Fig. 6.8 Vortices in a rotated superfluid

Coming back to superconductivity, the phenomenological theory was formulated by Vitaly Ginzburg and Lev Landau (1950), using Landau's magic wand once again, before the intrinsic mechanism became known. However, these equations are more complicated than those of the superfluid condensate, since besides the order parameter, they must also include the magnetic field, and to be compatible with electrodynamics, they have to be invariant under translations of the phase of the complex order parameter. All this was, of course, taken care of. It was also important that

Fig. 6.9 *Left*: Magnetic field avoiding a superconductor. *Right*: A levitating magnet

the Ginzburg–Landau equations contain the length scale that determined the penetration depth of the magnetic field into a superconducting material. Long before that, Walther Meissner and Robert Ochsenfeld (1933) had found experimentally that magnetic field does not penetrate into superconductors (see Fig. 6.9). This effect can be used to levitate a magnet over a superconducting material, and if refrigeration was cheap, maglev trains could be levitated in this way. The phenomenon was soon modeled by the brothers Fritz and Heinz London (1935), who suggested a simple equation for the magnetic field containing a small penetration length. The new theory explained this as well, but it had a new feature which even its creators failed to appreciate at first.

If the magnetic field is sufficiently strong, but not strong enough to destroy the superconductivity, the material separates into superconducting and normal domains, and the lines of magnetic field pass through the latter. The Ginzburg–Landau theory contains a parameter, call it μ, defining the surface tension of the boundary of these domains. We saw in Sect. 5.3 that the surface tension tends to make the boundary between different phases as short as possible at a given area or volume, and therefore domains should be large and have smooth boundaries. The young Alexei Abrikosov (1957) noticed that, if μ exceeds a certain critical value, the surface tension vanishes and then becomes *negative*. Negative surface tension means that non-superconducting domains would tend to disperse as finely as possible. This can be done by forming what came to be called Abrikosov vortices, similar to the vortices in superfluids in Fig. 6.8, oriented along the magnetic field and each carrying a quantum of phase circulation.

Abrikosov later reminisced that when he presented this idea to Landau, he was told that this was nonsense, because surface tension could not be negative – indeed, it never is in normal materials. Landau assumed things to behave in a generic way and blocked publication. This may well have happened, although Ginzburg, Abrikosov's co-recipient of the belated 2003 Nobel Prize in Physics, did not confirm it (the two did not really like each other). Landau could not be asked to confirm. Inconsiderate people said that he was the only one who ever got the Nobel Prize posthumously – in 1962, following the fatal car crash that elicited world-wide compassion and support; he lived for another six years but could no longer work. Landau, though not a fan of the Soviet regime (he did a stint in jail and only Kapitza's intercession saved him), could be as authoritative as Stalin in his field. The situation soon clarified, however. The first superconductors, mostly pure metals, did indeed behave in a way that seemed natural. They came to be called *type I* superconductors when *type II* superconductors, mostly alloys, were discovered. The new arrivals behaved in exactly the way Abrikosov had predicted and his paper was promptly published.

The microscopic BCS theory was finally presented by John Bardeen, Leon Cooper, and Robert Schrieffer (1957). The difficulty lay in the fact that electrons are fermions which, unlike bosonic particles, cannot condense into a single quantum state. The BCS theory asserted that electrons combine to form *Cooper pairs* which overcome the repulsion between negative charges through the attraction caused by their joint interaction with phonons – oscillations of the crystalline lattice (Fig. 6.10). Paired electrons are already bosons, and can form a superfluid condensate.

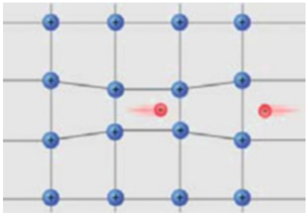


Fig. 6.10 Formation of Cooper pairs

Much more could be said about the continuing history of superconductivity. It is never perfect. The motion of vortices among impurities, hard both to model and to avoid, is a major cause of losses. On the other hand, the condensate turns out to be able to tunnel through a thin insulating layer between two superconducting layers. The discovery of this effect by Brian Josephson (1962), still a student, brought him the 1973 Nobel Prize in Physics, whereupon he immersed himself in transcendental meditation and paranormal phenomena, getting involved with the "Fundamental Fysiks Group" (Kaiser, 2011) mentioned in Sect. 6.3.

Georg Bednorz and Alex Müller discovered superconductivity at unusually high temperatures in ceramic materials (Bednorz and Müller, 1986), and were promptly awarded the 1987 Nobel Prize in Physics. The discovery generated frantic activity, and the transition temperature was soon raised above the boiling point of nitrogen, which makes refrigeration much cheaper. High-temperature superconductors, however, lack the stability necessary for their practical use, and the physical mechanism is still unknown. High in the skies, neutron stars are believed to be in a superconducting state. This would be an ultimate high-temperature superconductor, at millions of degrees kelvin, but hardly of any practical use for us here on Earth.

Going back to superfluidity, a rare isotope of helium, ^{3}He with only a single neutron in its nucleus, became available in considerable quantities as a product of the decay of the superheavy isotope of hydrogen ^{3}H used in H-bombs. ^{3}He, unlike common ^{4}He atoms, are fermions and need to be paired similarly to Cooper pairs to form a superfluid condensate at just a few thousandths of a degree above the absolute zero temperature (Osheroff et al, 1972); another Nobel Prize in Physics was shared in 1996 by the three discoverers. Paired ^{3}He atoms are far more sophisticated than electronic Cooper pairs. They are distinguished by spin and orientation; the condensate can exist in different superfluid phases and contain, not just one kind of vortex, but different kinds of topological defects (Volovik, 1992). An issue still debated by theorists is the existence of a supersolid state, in which a crystal flows like a liquid with zero viscosity.

Why would we be interested in these exotic states of matter? Imaginative science is driven by curiosity rather than by practical needs, but more than curiosity is involved here. The creation of defects in superfluid ^{3}He has been promulgated as a "cosmological experiment", gaining more insight at greater economy than anything dreamed of in gargantuan accelerators. In an earlier proposal (Zurek, 1985), even media with far simpler topology – either liquid crystals (Sect. 5.5) or ^{4}He – were supposed to play the role of a simulated vacuum. This may add body to field theory and spice to condensed matter studies: unlike ^{3}He, fascinating the experts because of the complex topology of its order parameter space, the creation of the Universe is an interesting topic in itself, and appeals to a much wider audience. However, the dynamics of defects in condensed matter, on the one hand, and defects with identical topology in relativistic spacetime, on the other hand, hardly have much in common.

Much more can be said of interactions between electrons, magnetic fields, and the structure of condensed matter. Their study involves sophisticated experimental and theoretical techniques inaccessible to outsiders. They may lack the fascination of other-worldly speculations, but they sometimes inspire and challenge the phenomenology of particle physics. The particles involved in quantum condensed matter are emergent quasiparticles, and this suggests that elementary particles might also be quasiparticles emerging from a deep structure of the vacuum on the Planck scale. Condensed matter physics may manipulate the number of dimensions – not increasing them, of course, but going down to two-dimensional phenomena that display some unusual features. In two dimensions, some quasiparticles, called *anyons*, obey fractional statistics, differing from both fermions and bosons. In topological insulators, conductivity is restricted to thin surface layers, giving a real-life example of the dimensional reduction that may restrict observable phenomena from whatever number of space dimensions to our four-dimensional space-time. In the 21st century, Nobel Prizes in Physics have equally honored studies in the physics of elementary particles and cosmology, on the one hand, and in condensed matter, on the other, and in the late 20th century the latter was given some preference. In neither of these discoveries was "quantum weirdness" an issue.

Chapter 7
Broken symmetry

7.1 Convective Patterns

When all basic laws are known, many surprises may still remain in their implementation. This sentence should be corrected: we may not know *all* basic laws, but we know those pertaining to a particular range of phenomena. In Sects. 6.4–6.6, we discussed the profusion of chemistry and quantum condensed matter hidden behind Dirac's bold assertion that a large part of physics and the whole of chemistry were already completely known. We can see it also on a simpler level. The Navier–Stokes equation of classical hydrodynamics describes the motion of Newtonian fluids perfectly, at least as long as we exclude singular short-scale effects, such as those near three-phase contact lines (Sect. 5.4). The equations of heat and mass transport (Sect. 3.2) are as reliable on the same macroscopic level. But does that mean we can predict weather? Of course not! The variety of flows and winds is enormous, and even simpler engineering problems challenge our most advanced computers. We shall come to the complexities of turbulence, but pause for a while on instabilities disturbing still waters.

One of them, surface waves, we already encountered in Sect. 2.5 – but these waves are caused either by a local perturbation, like throwing a pebble, or by wind; in each case the symmetry is broken by outside interference. The first experiment demonstrating *spontaneous* symmetry breaking was carried out by Michael Faraday (1831), who observed standing surface waves in a vertically oscillating fluid layer (a layer of sand can serve the same purpose). The basic scheme of experiments of this kind, repeated in other settings and not only in hydrodynamics, is shown in Fig. 7.1 (left). We start with a featureless plane. The external input is directed perpendicularly, without disturbing the symmetry – but the result is a *pattern*, as we see in the adjacent picture. Its wavelength depends both on the properties of the fluid layer – its thickness, density, and viscosity – and on the amplitude and frequency of the input. The dominant frequency of the surface oscillations is half that of the external input, a result that is not obvious.

L. Pismen, *The Swings of Science*,
https://doi.org/10.1007/978-3-319-99777-3_7

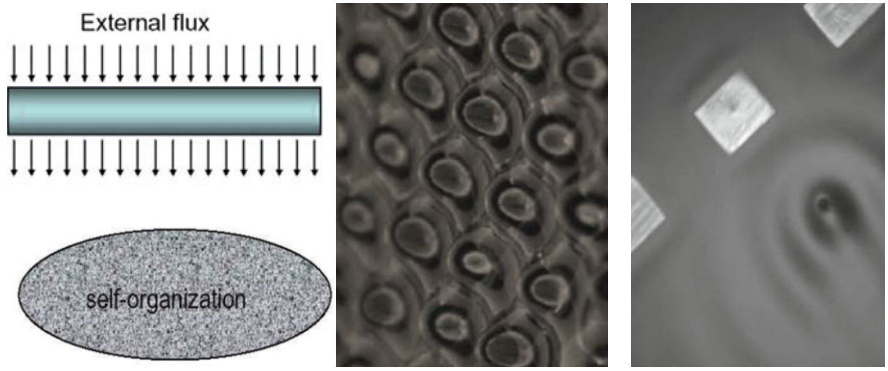


Fig. 7.1 *Left*: A scheme of pattern formation under the action of a uniform external input. *Center*: A hexagonal pattern of Faraday waves. *Right*: Faraday waves in the experiment imitating a quantum mechanical pilot wave passing through a double slit (Sect. 6.3)

The precursor of Faraday waves was the figures observed by Ernst Chladni (1787), patterns formed in a thin layer of sand covering a metal plate touched by a bow, picking out the nodal lines of resonant vibrations (Fig. 7.2).

The next experiment, the one that became most famous, was carried out by Henri Bénard (1900), who observed convection cells in a thin layer of whale oil heated from below. The mechanism, sketched in Fig. 7.3, was explained by Lord Rayleigh (1916): the light warm fluid rises upward, cools down there, and descends. Rayleigh carried out linear stability analysis and computed the critical temperature difference across the layer, above which convection starts. His work and the patriarch's authority were so impressive that the phenomenon came to be called Rayleigh–Bénard convection. This was fortunate in some way, because Rayleigh's

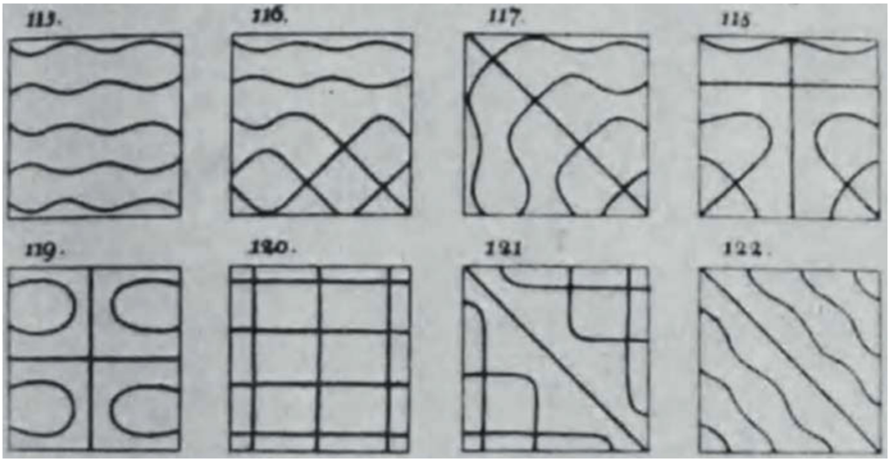


Fig. 7.2 Chladni figures

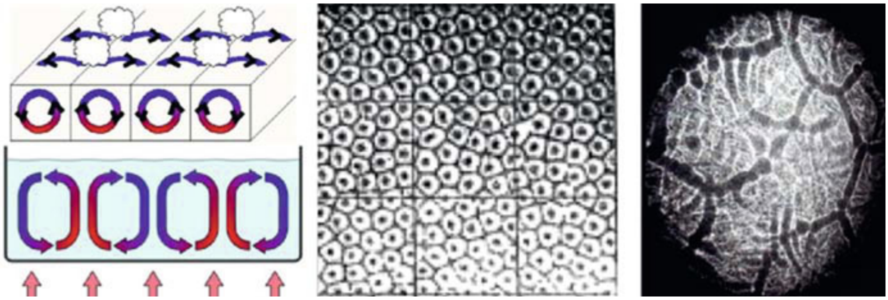


Fig. 7.3 *Left*: Bénard convection and the formation of cloud streets as ascending warm air cools and water vapor condenses. *Center*: A pattern of convection cells, in the original photo by Bénard (1900). *Right*: A snapshot of turbulent Bénard convection

explanation of Bénard's experiments turned out to be wrong! A later reproduction of the original experiment showed that the actual mechanism was quite different. The layer had a free upper surface, and instability was caused by the Marangoni effect (Sect. 5.3) driving the fluid along its surface from locations warmed up by ascending currents to cooler patches. The convection patterns are similar, and the name Bénard–Marangoni convection distinguishes it from the gravity driven phenomenon, which is more common, and is the only one possible when the fluid is confined between two solid plates.

Bénard convection has been thoroughly studied in precision experiments (Ahlers, 1974). It can generate not only hexagonal patterns, usually slightly distorted, as in the original experiments, but also patterns of different kinds, and it becomes turbulent when the temperature difference across the layer is high or the viscosity is low. A nonlinear theory of convective patterns has been developed in parallel to explain the observed phenomena in detail (Normand et al, 1977). Simulations of convective patterns imitate experiment, not only with regard to order, but also with regard to elements of disorder, like strings of defects separating hexagonal cells with different orientation or coexisting hexagonal and striped patches, as shown in Fig. 7.4. Con-

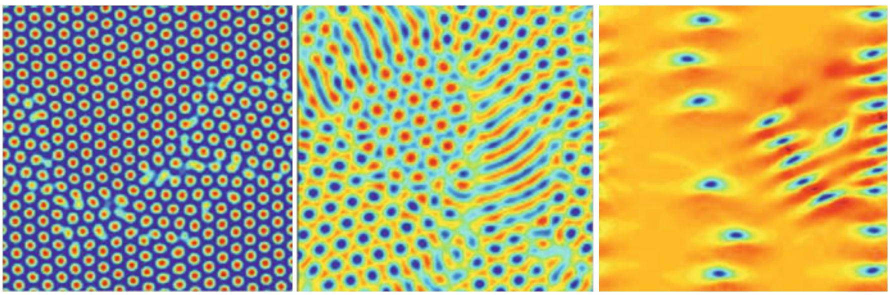


Fig. 7.4 Distorted patterns in simulations of Bénard–Marangoni convection with a deformable interface (Golovin et al, 2002)

vection driven by thermal and concentration gradients in conjunction with gravity plays an important role in industrial processes. The patterns are not so neat and regular as in specially designed experiments, but play a very important role in enhancing transport processes. It is also a crucial factor in meteorology, as air warmed up near the ground ascends, circulates, and drives cloud formation, as shown in the upper left panel of Fig. 7.3. Looking up into the sky or down from a plane, you can often see cloudlets forming nice parallel streets or almost regular hexagonal patterns.

An example of purely hydrodynamic pattern formation is Taylor vortex flow between two rotating cylinders. The basic state is not quiescent as in the above examples, but a parallel circular flow named after Maurice Couette, who used it as a well-ordered device for measuring viscosity. Geoffrey Taylor (1923) observed that, above a certain rotation frequency (or, to be precise, above a dimensionless combination of the rotation frequency, viscosity, and radii of the two cylinders, called the Taylor number), the parallel flow breaks into a chain of vortices surrounding the inner cylinder (Fig. 7.5). The flow is counter-rotating in alternating pairs of vortices. This pattern, in turn, exhibits a sequence of instabilities at higher Taylor numbers, when the vortices become wavy, modulated, and eventually turbulent. A similar instability takes place when the basic flow has an axial component, in addition to the circular one. It results in the formation of Taylor–Görtler vortices, a combination of Taylor vortices with Görtler vortices formed in a boundary layer at a concave wall; an example is shown in Fig. 7.5 (right). This flow pattern has been studied less extensively, but it is also apt to exhibit a sequence of secondary instabilities.

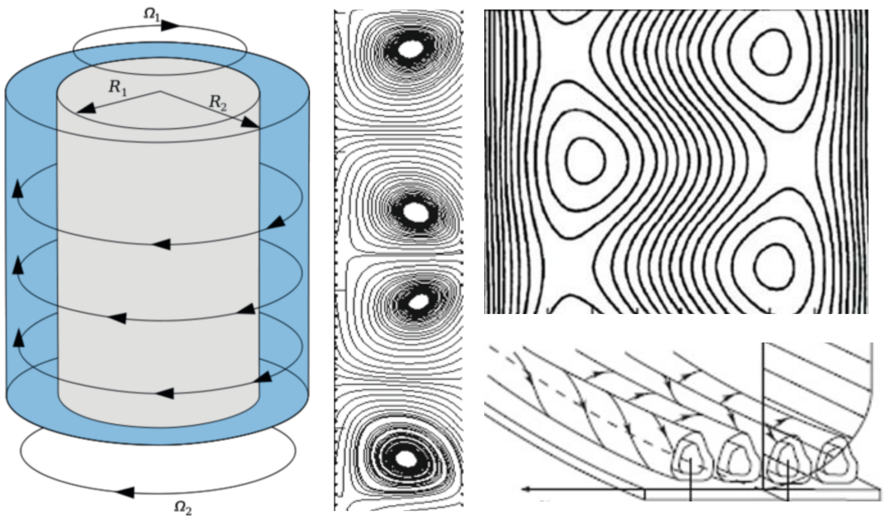


Fig. 7.5 *Left*: Apparatus for studying Taylor vortices. *Center*: Taylor vortices. *Right*: Taylor–Görtler vortices (*upper*) and Görtler vortices (*lower*)

7.2 Chemical Patterns

In chemical and biological applications, awareness of the instabilities causing spontaneous pattern formation had to wait another half a century for the famous work by Alan Turing (1952), far weaker analytically than that of Rayleigh, but philosophically charged and winged by the fame of the Turing machine and the Enigma Code. The paper bears the ambitious title "The chemical basis of morphogenesis" but ends on a humble note: "It must be admitted that the biological examples which it has been possible to give in the present paper are very limited. This can be ascribed quite simply to the fact that biological phenomena are usually very complicated." It is actually not about biology, but about chemical patterns, and the rational message to be extracted from the 36 long pages is that pattern formation requires, in the simplest setting, the combination of a slowly diffusing activator and a rapidly diffusing inhibitor. This principle, which can be established in a few lines by the linear stability analysis of a two-component reaction–diffusion system, is prominent in many model pattern-forming systems. Both activator and inhibitor can be either chemical or biological species, or, in a more abstract form, other physical agents, and the scale of the pattern is determined by the relevant spreading ranges.

The way a Turing pattern is formed is schematized in Fig. 7.6 (left). If the level of the activator is raised locally (upper panel), it also raises the level of the inhibitor (middle panel). Since the latter is more diffusive, it spreads out to an area where the activator level is low, depresses the activator there, and as the activator level goes down, is itself depressed (lower panel). A biological analogy would have herbivorous animals attracted by the local abundance of grass and trampling it underfoot. The pattern of higher and lower levels of the two species propagates out with the wavelength about a geometric mean of the diffusion ranges of the two species. The final result may depend on the geometry of the region, on some special interactions, or just on initial conditions. Typical patterns are slightly distorted hexagonal, as in the central panel of Fig. 7.6, or striped, turning into labyrinthine ones due to the presence of dislocations, as in the right-hand panel. Patterns may evolve with time

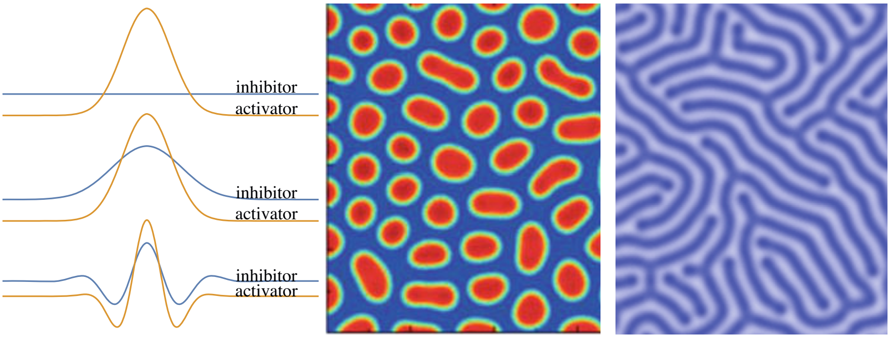


Fig. 7.6 *Left*: Patterning involving a slowly diffusing activator with a rapidly diffusing inhibitor. *Center*: Hexagonal pattern. *Right*: Labyrinthine pattern

into more regular shapes, e.g., as dislocations collide and annihilate, but this process, similar to coarsening (Sect. 5.3), is slow and can be frustrated by the various inhomogeneities. The same model has been successfully applied to desert vegetation patterns (Meron, 2018), where plants served as an activator and the *lack* of ground water sucked in by plants as an inhibitor.

Chemical patterns of another kind are *dynamic*. In hydrodynamics, all kinds of dynamic wave patterns are typical (Sect. 2.5), while stationary patterns of the kind described in Sect. 7.1 (where, of course, the fluid still moves within cells or vortices) need special arrangements. We shall see that the situation with chemical patterns is actually similar – but chemical oscillations were considered taboo, because of thermodynamical misconceptions, well into the 1960s, in spite of the fact that electrochemical oscillations and waves have been known since the turn of the 20th century (Ostwald, 1900). Turing (1952) did not consider this possibility either. Chemical oscillations were discovered in 1951 by Boris Belousov, but the prejudice was so great that he was barred from publishing it, and only in 1959 managed to squeeze a note into an obscure collection of medical abstracts.

Moscow has always lived on rumors more than on publications, and the discovery did not go unnoticed. A young graduate student Anatol Zhabotinsky (a grand-nephew of the famous Revisionist Zionist) was encouraged by his supervisor to try and find out how Belousov's recipe worked. It was a success. Zhabotinsky (1964) went on to observe fascinating wave patterns in a Petri dish. The entire story was later described by Winfree (1984). By this time, what came to be called the Belousov–Zhabotinsky (BZ) reaction was high fashion. Belousov was honored posthumously by the Lenin Prize, together with his much younger Soviet followers. Irving Epstein, the "nonlinear chemist" who was Zhabotinsky's host at Brandeis University in the last 16 years of his life, mentioned in Zhabotinsky's obituary that Ilya Prigogine "regarded the BZ reaction as the most important scientific discovery of the 20th century, surpassing quantum theory and relativity". This exaggeration may be taken as another face of the old prejudice, but Ilya Prigogine, though called by some critics the least deserving of the Nobel scientists, was more sophisticated. Perhaps he was referring to a connection with life whose unending oscillations might overcome the dreaded "arrow of time"? It is too late to ask. Of course, BZ oscillations end when the reactant is exhausted, as do our lives.

I recall Zhabotinsky giving a talk at Levich's theory department at the time when the fame of the BZ reaction had not yet spread. The group studying electrochemistry of nerve impulses and myself, already fixed on symmetry breaking, talked with him about a theory of his target and spiral waves. In retrospect, it was so easy! Why didn't we do it right away? But what is a theory? The exact chemical mechanism is disputed to this day, but who cares besides the chemists involved? Other oscillating chemical reactions are now known. Patterns and waves on a much finer scale, and with added anisotropy distorting round spirals to a squared form, were observed on catalytic surfaces, earning Gerhard Ertl the 2007 Nobel Prize in Chemistry. Their mechanism, involving the restructuring of the catalytic surface (Fig. 4.7), is still more complex. What is easy is a qualitative explanation, similar to Turing's recipe, and this is what we failed to understand, although a suitably simple system was al-

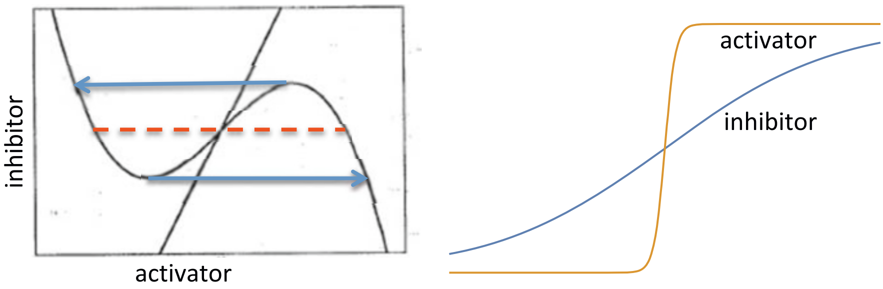


Fig. 7.7 *Left*: Activator and inhibitor dynamics in the FN system. *Right*: A stationary activator/inhibitor front

ready waiting to be applied. This was the FitzHugh–Nagumo (FN) equation, a caricature of a caricature, first derived as a simplified version of the Hodgkin–Huxley nerve conduction model (FitzHugh, 1961). In another breakthrough for simplicity, a basic combination of two reaction–diffusion equations, first a nonlinear activator equation, and second an inhibitor equation, which can be linear, can generate the various patterns and waves in different applications, both in chemistry and biology.

The way this system works is illustrated in the left-hand panel of Fig. 7.7. The net formation rate of the activator vanishes on the S-shaped curve in the plot spanned by the activator and inhibitor concentrations, while the inhibitor concentration grows to the right and decreases to the left of the inclined straight line. The only stationary state is the intersection of the straight and S-shaped lines in the center, and it is unstable. If the activator level lies on the left-hand branch of the S-shaped curve, the inhibitor decays until its level reaches the bottom of the S-curve. Beyond this point, a stationary activator level cannot be sustained, and it rises fast to reach the right-hand branch. Now the inhibitor concentration goes up until it reaches the upper bend of the S-curve, the activator level drops back onto the left-hand branch, and so it goes. This temporal oscillation translates into a wave pattern when its phase changes from place to place. The two inhibitor levels indicated by the arrows in the picture are then translated into the front and back of the activator wave.

The same system describes stationary fronts separating alternative domains in Fig. 7.6. The sharp activator front in the right-hand panel of Fig. 7.7 is stationary when the inhibitor concentration at its location rapidly relaxes to the level shown by the horizontal line in the left-hand panel, such that the two areas between this line and the S-curve are equal. When the inhibitor is not as rapid, stationary fronts are destabilized and start moving. A wider diffusion range of the inhibitor responsible for a gentle slope in its spatial distribution compared to the sharp activator front is necessary for the formation of stationary (Turing) patterns, but is irrelevant for oscillations and waves. The latter are more natural; a graduate student with no funds armed only with an old chemist's recipe could see them in a Petri dish, while stationary patterns required a careful experimental setup, with a uniform steady supply of reactants and a liquid solution replaced by a gel to suppress convection, and they

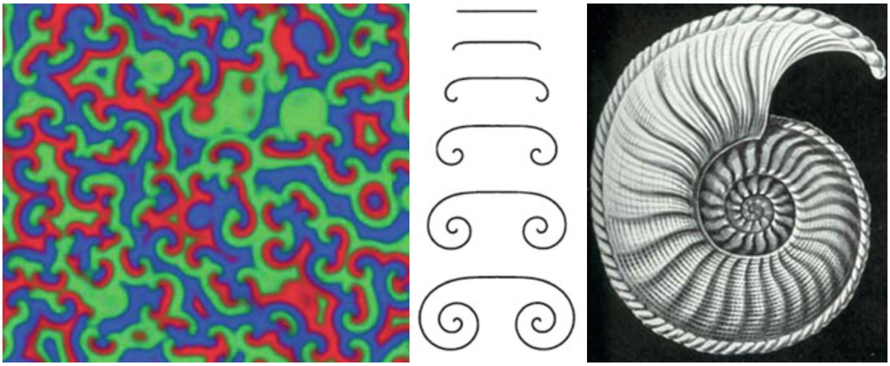


Fig. 7.8 *Left*: Snapshot of the spiral wave pattern of the BZ reaction. *Center*: Formation of a spiral wave. *Right*: Spiral shape of a fossil ammonite (Haeckel, 1904)

were only observed in the same BZ reaction many years later (Ouyang and Swinney, 1991).

It is also not so difficult to see why reaction fronts coil into spirals, as in Fig. 7.8. What Zhabotinsky first observed were target patterns, but if a circular front was disturbed by any obstacle (just inserting a finger would do), it turned into a spiral. The central panel of Fig. 7.8 illustrates this. The edge of a front segment propagating upwards in the picture lags behind and, as the segment spreads out, spirals are formed at its sides. Spirals are ubiquitous in nature but are usually formed, as in the fossil conch in the right-hand panel, by peculiarities of the growth process, rather than symmetry-breaking instabilities.

7.3 Unity in Variety

In his later years, Gottfried Wilhelm Leibniz (1714) pondered the idea that God, being absolutely perfect, would create the best world that God could, and being omnipotent and omniscient, would create the best of all possible worlds. Voltaire lampooned Leibniz as the wretched Dr. Pangloss in *Candide*, who maintains in the middle of every disaster that we are living in the best of all worlds. Leibniz's contemplations were the continuation of age-long efforts to deny what is obvious to everybody struggling with evil, and his notion that God is metaphysically necessary contradicts the logical impossibility of God existing outside the world and having nothing to care about. But the following statement fits our current subject well: "to get the greatest possible variety, but with all the order there could be, i.e., it is the way to get as much perfection as there could be" (his # 58). Symmetry breaking is the most straightforward way to generate variety, though it is apt to generate it in both good and evil ways. We would not exist in the world without variety, and therefore variety is good for us, even if it sometimes turns out not to be to our taste.

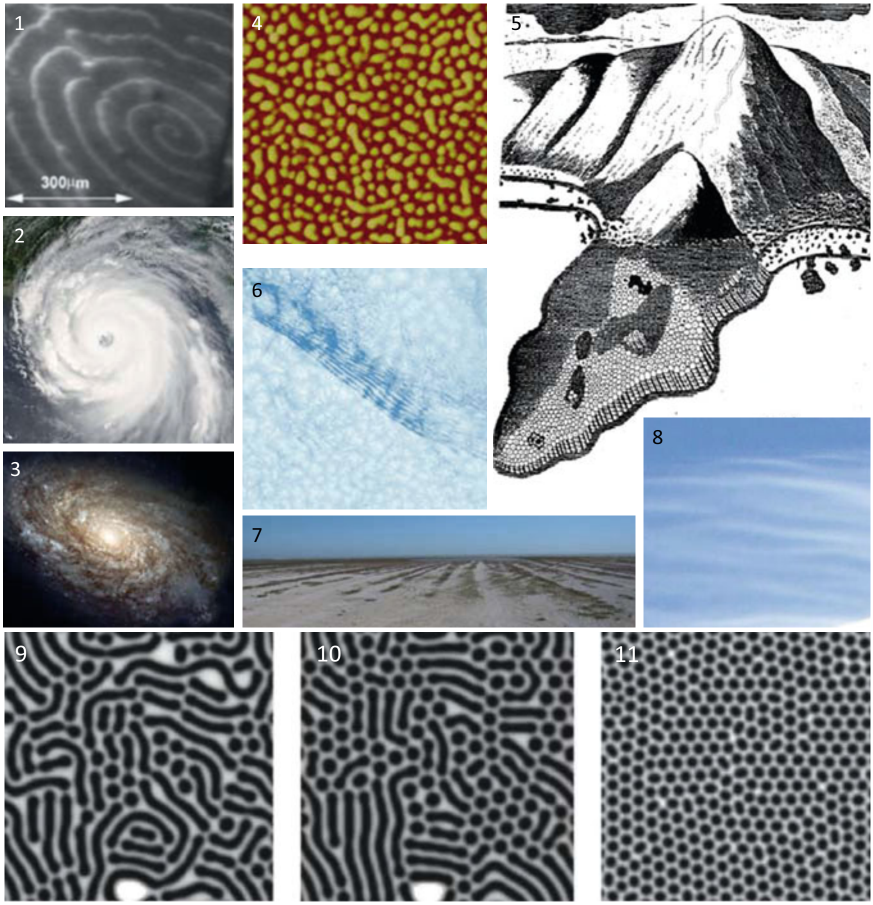


Fig. 7.9 (1–3): Spiral patterns. (1) Oscillations on a catalytic surface. (2) Hurricane Katrina at peak intensity. (3) A spiral galaxy (NASA image). (4–6): Hexagonal patterns. (4) Separation of the components in a copolymer. (5) The Devil's Causeway in Ireland (a 1694 drawing). (6) A cloud pattern (NASA image). (7–8): Striped patterns. (7) Shingle ridges on the Suffolk coast. (8) A cloud pattern (author's photo). (9–11): Transition from disordered strips to hexagons in a dynamic phase separation model (Golovin and Pismen, 2004)

Symmetry breaking works in qualitatively similar ways under different circumstances and in completely different physical systems. Even external fluxes maintaining the system out of equilibrium as in Fig. 7.1 are not required. A helpful example is *copolymerization*, binding two mutually repelling components into the same chain. This brings up the issue of social associations, but we will talk about that later (Sect. 10.2). If the two monomer species were free, they would separate into two phases, but it is impossible when they are bound in a polymer chain. They try their best, however, and cluster in separate neighborhoods that are as large as can be allowed by the freedom of bending flexible macromolecular chains. The result is a pattern looking like any other, as shown in panel 4 of Fig. 7.9. Another example of this kind is the deformation of a thin crust over a drying soft layer, yielding a visually indistinguishable pattern, even though driven by a quite different mechanism. Solidification is also a symmetry-breaking transition on the molecular scale, from a featureless liquid to an ordered crystal.

Crystalline forms are plentiful, and the same system – be it convection, vibration, or chemical reactions – is quite often able to generate different patterns. On the other hand, different systems may generate the same pattern, so that we would not recognize its origin from a picture. Both convective and chemical patterns may look similar to desert vegetation patterns or to the patterns of animal coats. Convective patterns with a similar structure can be generated on characteristic scales from less than a millimeter in electro-convection in liquid crystals to kilometers in ocean currents. Strong inhomogeneities can develop on the scale of nanometers in copolymers and on the scale of millions of light-years in galactic clusters. Spirals are formed for entirely different reasons and on immensely different scales in chemical oscillations, growing plants, hurricane cloud formations, and spiral galaxies. A collection of pictures is brought together in Fig. 7.9. What is the origin of the hexagonal tiling of the famous Devil's Causeway, a pattern also seen in other parts of the world, like the Hexagon Pool in the Golan Heights? Here, we are very likely seeing solidified Bénard cells of convection in lava – but crack patterns may also assume a roughly hexagonal form. This collection could be greatly expanded, including also desert vegetation patterns and patterns in nonlinear optical circuits.

When physicists encounter such a situation, with similar basic structures arising in physically unrelated systems, often driven by complicated and only vaguely understood mechanisms, what is their reaction? Leave it to engineers, geologists, and biologists to rack their brains? The physicist's way to meet the challenge is to invent generic models replicating observed patterns. The FN system is an *ad hoc* model of this kind, but a more rational approach is based on symmetry considerations and economy of tools. The origin of some model equations is similar to Landau's theory of phase transitions. One model of this kind is the Swift–Hohenberg (SH) equation (Swift and Hohenberg, 1977). While the standard Landau theory of phase transitions (Sect. 5.2) leads to what is called the Ginzburg–Landau equation with the negative cubic term limiting growth and the diffusion term encouraging phase separation, the SH model inverts the sign of the diffusion term. This would cause a catastrophic evolution, breaking the coexisting phases into tiny pieces, but it is prevented by a higher-order term containing the fourth spatial derivative. The equation is dissipative and evolves to an equilibrium state which will usually be a striped pattern, where the width of the stripes is regulated by the parameters of this equation.

This sounds bland, but playing with parameters, we can generate more interesting situations when two separate uniform states coexist with a striped pattern. This coexistence can be disrupted by a sufficiently strong perturbation. What happens is shown in the two upper panels on the left of Fig. 7.10. Note that in both cases the striped pattern pervading the entire domain is distorted in different ways. It is also distorted, and also differently, when it emerges from an absolutely unstable uniform state following a weak random perturbation. Evolution to a well ordered pattern

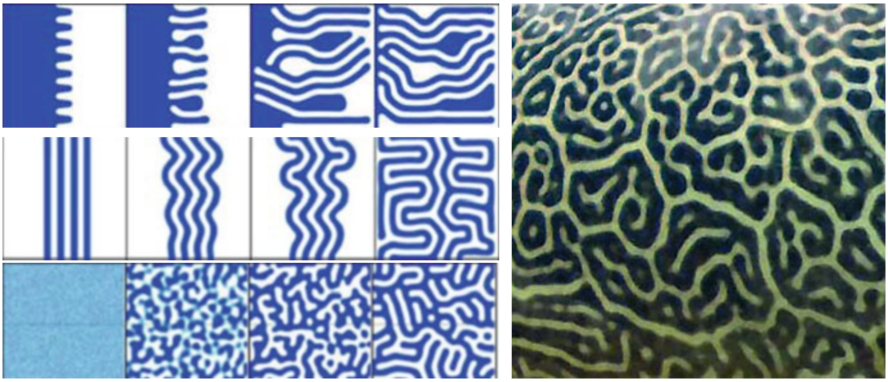


Fig. 7.10 *Left panels*: Numerical solutions of the SH equation (Hagberg et al, 2006). *Upper*: A strong perturbation of the boundary between the two coexisting phases (dark and white) leads to the development of an irregular striped pattern. *Center*: A striped pattern coexisting with a uniform state invades it when perturbed. *Lower*: A random perturbation of an unstable uniform state develops into a labyrinthine pattern. *Right*: A fish skin pattern

is very slow, like all ripening processes (Sect. 5.3). The pattern in the right-hand panel of Fig. 7.6 also looks labyrinthine, and so does the pattern of a fish skin in the right-hand panel of Fig. 7.10, but this does not mean that it has been formed by a mechanism obeying either the SH or FN equation. The simulations of realistic-looking animal skins by model equations appearing in some publications should not be taken seriously.

In order to obtain stationary hexagonal patterns similar to those appearing in a number of pictures in this chapter, the inversion symmetry of the SH equation must be broken by adding a quadratic term. Let us take a look at any asymmetric nonlinear system, not just the modified SH, which becomes unstable under perturbations with a certain wavelength, or which comes to the same, under perturbing Fourier modes with a certain wavenumber (inverse wavelength) and wave vectors which, because of the spatial symmetry, are arbitrarily directed. If we start its evolution from a perturbed unstable homogeneous state, the quadratic term induces a resonance between wave vectors forming an equilateral triangle, and these combine into a hexagonal pattern in the plane. In three dimensions, a greater number of resonant modes will be excited, forming the edges of a tetrahedron, an octahedron, or an icosahedron. The former two correspond to the face-centered and body-centered cubic crystal lattices, and the latter to a quasicrystal.

I was long fascinated by a possible connection between these triple resonances and quark triplets tied up in hadrons. A triple resonance can involve only space-like modes in spacetime – tachyons traveling faster than light. I imagined hadrons as long-scale time-like envelopes of resonant triangles on the Planck scale, and from time to time put regular doings aside and tried to work out bifurcations that would lead to realistic results in this way or another until the attempt collapsed. I imagined these regular tetrahedra, which cannot be closely packed in the flat space, to be packed when the space is strongly curved, and the compact spatial structure corresponding to this polyhedron exploding at the Big Bang. I never told anything of this to anybody, but maybe I was not alone among my colleagues to engage in such fantasies. The fascination of "nonlinear science" lies in its inherent similarity to particle physics, minus spacetime, minus Nobel Prizes, plus complex phenomena caused by dissipation and external pumping of energy.

7.4 Instabilities That Never Saturate

Besides stationary or dynamic patterns, another common structure seen in various non-equilibrium systems from phase transitions to living forms, is *dendritic*. This structure is typically generated by growth. If a crystal grows by accretion of atoms diffusing from a solution, its flat surface is unstable. Suppose there is slight protuberance on the surface: it is better exposed to the solution and grows further. In contrast, a dimple is less accessible and is left further behind. The farther a protuberance grows, the closer it comes to the source of material, its sides lose stability, and it starts to branch out in its turn. The result is a dendritic structure like the one in Fig. 7.11 (top left).

Another way a dendrite may develop is through the Saffman–Taylor instability. It develops in a Hele-Shaw cell, a shallow space between two parallel plates where a liquid can be pumped either from the sides or from above, also drawn in Fig. 7.11 (bottom center). If a less viscous fluid displaces a more viscous one, a protuberance on the interface encounters less resistance and grows further, creating a fingering pattern like the one in Fig. 7.11 (bottom left). The upper central panel in this fig-

Fig. 7.11 *Upper left*: Dendritic copper crystals. *Lower left*: A dendritic flow pattern in a Hele-Shaw cell. *Upper center*: High voltage dielectric breakdown in a polymer block. *Lower center*: A Hele-Shaw cell. *Right*: Mycetozoa (Haeckel, 1904)

ure shows a branching electric discharge in a dielectric; sometimes we see forked lightning choosing several channels of discharge. The right-hand panel shows the various branched shapes of slime mold.

This instability never saturates, the interphase boundary remains unstable, and the dendrite keeps branching and growing. Of course, we are speaking of an ideal dendrite in an infinite medium. In reality, any instability saturates at some level that is not accounted for in the basic model. The growing crystal will slow down when its tip becomes too sharp, close to a molecular scale. A Saffman–Taylor finger is slowed down and kept from further branching by surface tension – but before this limit is reached, a highly branched pattern will develop.

The ultimate example of a catastrophic development is given by the Rayleigh–Taylor instability (Fig. 7.12). It starts from an absolutely unstable state of a fluid layer in an overturned position, say, a flat water film on the ceiling or any heavier fluid on top of a lighter one. Once the flat interface is perturbed, the instability develops catastrophically, as shown in the snapshots of a simulation, with ever finer structures formed (until, of course, the heavy liquid reaches the bottom and everything calms down).

What makes this instability so dramatic? It is the prospect of the Most Awful Instability of All that Ends the World. Our vacuum cannot be unstable like a water film on the ceiling, but it might well be metastable – it is then said to be a *false* vacuum, so that a sufficiently large perturbation would nucleate a region of *true* vacuum spreading with the speed of light and destroying everything in its way. I mentioned (Sect. 2.2) fears that CERN's monster might perturb our vacuum strongly enough – but stronger perturbations that could reach us fast enough do happen in faraway stars, and the Universe has still not fallen apart. Espinosa et al (2018) argue

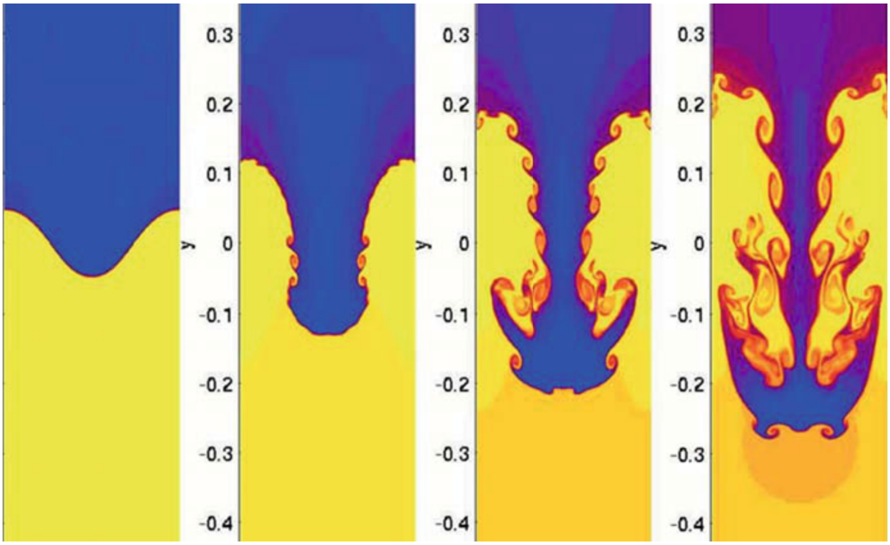


Fig. 7.12 Simulation of the development of the Rayleigh–Taylor instability

that such an instability might have happened during the presumed inflation stage following the Big Bang, seeding microscopic black holes. It would be reasonable to view the Big Bang itself as an instability of this kind, and the inflation stage as spreading our vacuum within another, metastable one, which we know nothing about. You can imagine Fig. 7.12 as a metaphor of a cosmic instability, and its whirls curling on ever finer scales (inaccessible to the simulation) as precursors of black holes. Espinosa et al consider these tiny but massive black holes as the essential component of dark matter that keeps our Universe from expanding too fast before emitting its ultimate whimper under the dark skies (Sect. 1.4).

Chapter 8
Complexity Simplified

8.1 From Order to Chaos

Our world might never have been more stable, both ontologically and politically, than at the time when it stood on the firm ground of the laws of Newtonian mechanics, developed into a beautiful mathematical structure by the brilliant French and German mathematicians of the Age of Enlightenment (Sect. 2.1). The social and moral ground was firm then as well, before being shaken by revolutions. Deterministic laws would, in principle, allow computation of all future states of a system, and Laplace's demon (Fig. 8.1) should have been able to accomplish this task:

> An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes. (Laplace, 1802).

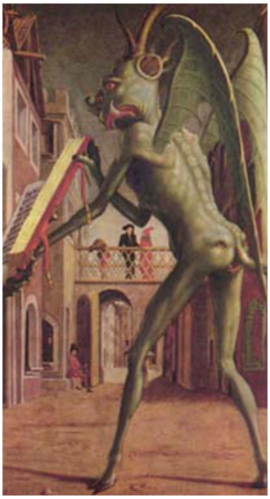


Fig. 8.1 Laplace's demon

At the birth of quantum mechanics, some physicists enjoyed and some deplored its probabilistic character. It was really a misplaced concern. Everybody cites, and I myself have already done so, Einstein's complaint that God does not play dice. But what is playing dice, or any chance game in an old-fashioned casino? It is a purely mechanical action perfectly obeying Newton's laws. Laplace's demon should definitely be able to break a casino. The point is to know all the initial conditions, "all positions of all items of which nature is composed", to infinite precision, a task that is truly demonic or divine or just impossible. Even with no quantum uncertainty, the future is never certain. Henri Poincaré (1890) became the first to express this

L. Pismen, *The Swings of Science*,
https://doi.org/10.1007/978-3-319-99777-3_8

uncertainty, known to everybody, in a rational form. He took part in, and won, the challenge announced by the King of Sweden to solve the three-body gravitation problem. After first mistakenly arriving at an apparently rational solution, he understood that a three-body system behaves chaotically. Vladimir Arnold (1964) proved that the Solar System is intrinsically chaotic. It is rendered unstable by "Arnold diffusion" due to weak multibody interactions, although thankfully on exceedingly long time scales.

Jacques Hadamard (1898) found that trajectories of frictionless particles on a billiard table with a concave wall ("Hadamard's billiard") diverge exponentially from one another, which is a characteristic of chaotic motion. The discovery of *deterministic chaos* by Poincaré and Hadamard was, however, not followed up for another half a century, while physicists were brooding about quantum uncertainty. The next chaotic system, also failing to attract much attention, was Khaikin's "universal circuit" devised in the course of studies of oscillations in electrical networks (Andronov et al, 1937). Chaotic behavior was also detected by Mary Cartwright and John Littlewood in the dynamics of the forced van der Pol oscillator, another model inspired by electric circuits (Cartwright and Littlewood, 1945). Chaos, however, did not become popular before the last third of the 20th century, when it prompted scientists and philosophers to rethink the fundamental questions of determinism, predictability, and randomness.

It had always been common knowledge that complex phenomena, influenced by a great number of different factors, were unpredictable (at least, from a practical point of view) and were amenable only to statistical methods. However, statistics gives little consolation when one has to live through a single decisive event, as gamblers, market traders, and weather forecasters know only too well. Predicting the future has always been a preoccupation of people in different walks of life, from clairvoyants to investment bankers, using no less diverse methods. It was therefore a kind of culture shock (at least for those who were able to comprehend it) that even very simple systems could sometimes behave in highly unpredictable ways.

8.2 Toy Models

The utterly simple chaotic models that helped chaos gain attention in the chaotic 1960s were discrete maps rather than differential equations. Bakers have always known perfectly well how to mix their dough. They stretch it to twice the length and fold it upon itself, so that, while the area is preserved, points that were initially far apart come together and points initially close become further separated. Mathematicians got hold of this idea and called it the "baker's transformation". If, instead of bending, you cut the dough and superimpose the pieces without flipping the direction, the transformation becomes the "Bernoulli shift". Why Bernoulli? Nicolas Bernoulli invented the Saint Petersburg lottery, where the initial stake is doubled every time heads appears, while in 1738 his more famous cousin Daniel, a long-standing resident of this city, resolved the paradox of the apparently infinite ex-

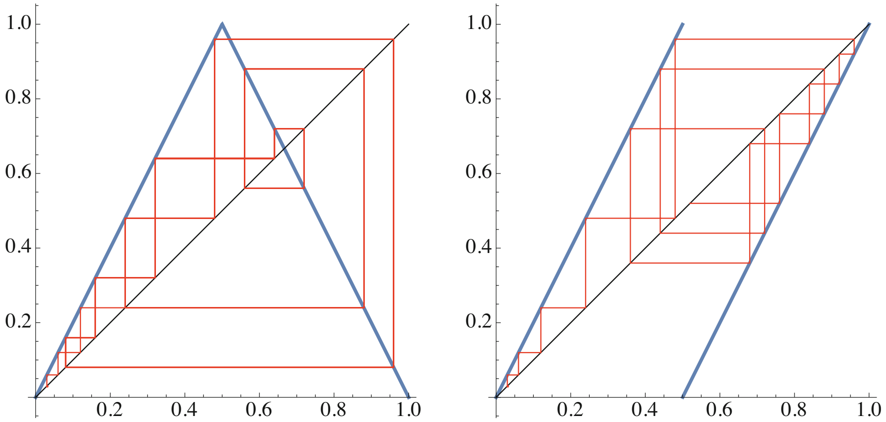


Fig. 8.2 Maps of the baker's transformation (*left*) and Bernoulli shift (*right*). *Thin lines* show successive iterations

pected yield of this lottery. If the position of a point on the unit interval is expressed as an infinite binary string, the leading digit 0 or 1 would mean, respectively, that it is located in either the first or the second half of the interval. After the Bernoulli shift, the second digit of the original string will classify the position of the point on the cut-and-moved piece; it will be the third digit after the next cut-and-move, and so on, which is equivalent to a repeated tossing of a fair coin. It is clear that, however close two points may be originally, they will eventually be separated and sent to the two alternative half-intervals: if the difference in the original positions was in the nth digit, this will happen after the nth iteration. The two maps in Fig. 8.2 show successive positions of a point under the repeated baker's transformation or Bernoulli shift.

In both the above maps, the total area, which plays here the role of what is called the "phase volume" in the theory of dynamical systems, is conserved. Many systems we encounter in real life are dissipative, and their phase volume shrinks with time. We could think it unlikely that, when the total area decreases, neighboring points can separate as they should if the dynamics is chaotic. The phase volume may, however, stretch in a certain direction while shrinking in others. As the original phase volume shrinks, it eventually converges to some *attractor*, and if the process is chaotic, this attractor should have a *fractal* structure. We met a fractal object in Fig. 1.1, at the very beginning of this book, where it served just as a nice illustration of a complex structure constructed in a simple way[1]. A fractal object gets this name because it has a fractal dimension.

We have no doubts about the dimension of the world we are living in, multidimensional worlds of string theorists excluded, and even there the dimension is always an integer, but mathematicians have invented different definitions for a frac-

[1] Benoit Mandelbrot (1982), who coined the term "fractal", popularized the Julia set and appropriated the name.

Fig. 8.3 Measuring the length of the coast of Britain using shorter and shorter rulers

tal dimension. They differ in the way they are computed and sometimes also in the result, but the main feature is that a fractal object has a higher dimension than its "topological" dimension (Farmer et al, 1983). Without the benefit of theory, Lewis Fry Richardson (1961) devised a simple algorithm while searching for a spurious relation between the probability of two countries going to war and the length of their common border. If you measure the length of a convoluted line, like the coast of Britain in Fig. 8.3, using a ruler and placing it in such a way that both its ends lie on the line, the total length increases as you decrease the length of the ruler, and it scales with a power equal to the line's fractal dimension. This is as good a way as any to define it. The dimension of the boundary of the Julia map approaches two – the dimension of a plane rather than that of a line.

Stephen Smale (1967) devised his "horseshoe map" to illustrate this mechanism in a simple way. The unit square in Fig. 8.4 is stretched horizontally and shrunk vertically in such ratios that its total area decreases (a). It is then folded and mapped back in such a way that the bent area remains outside the square (b). The part of the original square that is mapped back onto it following the transformations is restricted to two vertical strips (c). The area that is mapped upon itself is confined to the four squares (d), the intersections of the stripes in (b) and (c). The next iteration restricts the invariant area to the 16 small squares (e). Upon successive iterations, the invariant area shrinks further, all but fading from view, and eventually becomes a fractal object of zero area, but not of zero dimension as a finite number of points would be. Neighboring points of the original square can be found at far-removed locations, and predicting the location of their image after a large number of iterations would require us to define the initial data with ever-increasing accuracy.

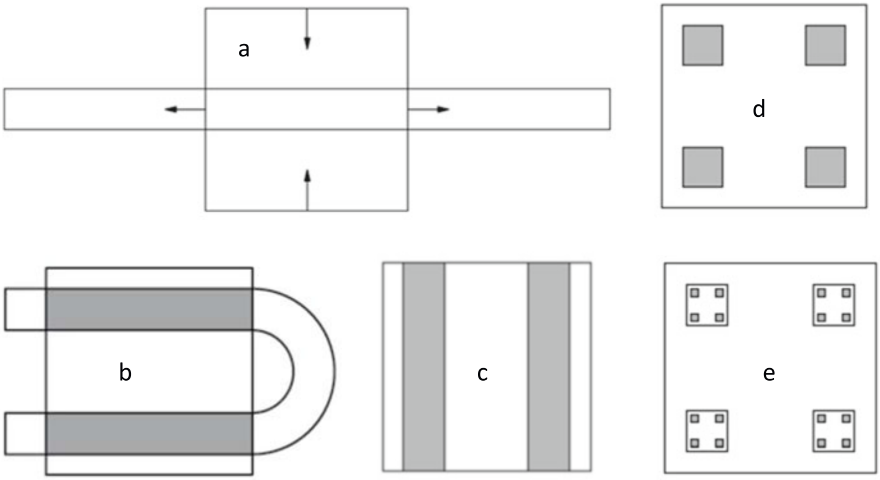


Fig. 8.4 Smale horseshoe. See the text for explanation

The most popular toy model is the *logistic map*. It is more realistic because it is continuous and does not involve geometrical rearrangements that are difficult to reproduce in common physical problems. I tried hard to avoid formulas but this one is very simple: take a point x_n in the interval between 0 and 1 and map it into the point $x_{n+1} = f(x_n)$, with the quadratic function $f(x) = \mu x_n(1-x_n)$ which ensures that the point never leaves the unit interval when the transformation is iterated as many times as we wish. If the parameter μ is small, iterations converge to some point; as μ grows it will start to alternate between two points at $\mu = 3$, then this number doubles at $\mu \approx 3.4495$, and keeps doubling further with ever decreasing parametric intervals, until the attractor becomes fractal and the dynamics becomes chaotic at $\mu \approx 3.569\,945\,6$. This period-doubling road to chaos was discovered by Robert May (1976).

As a theoretical ecologist, May was accustomed to smooth mappings that were used to relate populations in successive years, but he realized that continuous dynamics described by differential equations may also go along this road and become chaotic. Discrete maps are much simpler. They can be just one-dimensional, like the baker's transformation, the Bernoulli shift, and the logistic map. For continuous systems, the minimal dimension is three – otherwise trajectories, which are not allowed to intersect, cannot weave complex chaotic structures. There is, however, a way to convert a continuous system to a discrete one: the Poincaré map. The left-hand panel of Fig. 8.5 shows a trajectory of one of the three-variable equation systems devised *ad hoc* by Otto Rössler (1976) to look similar to the kinetic equations of chemical reactions. A Poincaré map is obtained by choosing a Poincaré section – a plane placed in such a way that the trajectory repeatedly crosses it, as in the central panel, and then plotting the intersection points, as in the right-hand panel of Fig. 8.5. Since the system is dissipative, its phase volume shrinks, as in Smale's horseshoe, and all

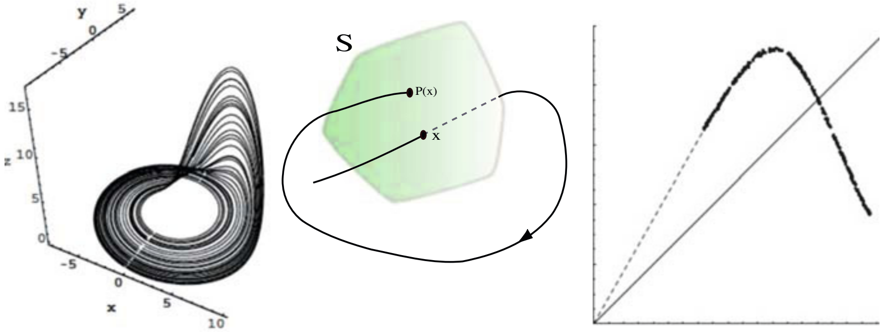


Fig. 8.5 *Left*: Trajectories of Rössler's system. *Center*: A point x is mapped onto $P(x)$ on the Poincaré section S. *Right*: Poincaré map of Rössler's system

points collapse to a single line that looks very much like the logistic map. This is a good indication that this system also behaves qualitatively in the same way as it becomes chaotic.

Mitchell Feigenbaum (1978) was exploring the logistic map helped only by a pocket calculator, a now forgotten contraption that was widespread at the time. Luckily for him, he did not have a faster device and had to think about how he could speed up the calculations. Iterated maps look very similar to each other if one cuts a piece near the point where the period doubles and rescales them. This suggested that some kind of renormalization, similar to the Nobel prizewinning theory of near-critical phenomena by Kenneth Wilson (Sect. 5.2), but much more easily implemented, might possibly be of some help. This was his life's coup. Period-doubling was viewed for some time as a universal way to chaos – but with time, as usual, complexity crept back again, and scenarios involving a transition to chaos turned out to be too numerous to be universal.

The chaotic attractor of the logistic map undergoes an infinite number of bifurcations as the parameter μ increases beyond the critical point. A coarse picture like the one in Fig. 8.6 still gives some idea of the true complexity of behavior that would be

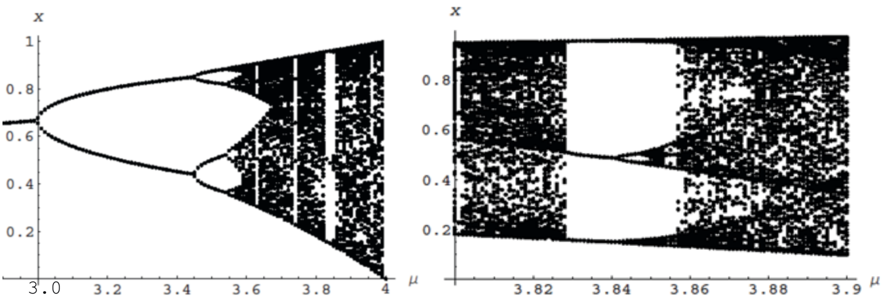


Fig. 8.6 Points on the attractor at different values of μ

revealed at finer scales when it is enlarged. The sequence contains an infinite number of periodic "windows" and repeated cascades. A prominent feature is a rather wide period-three window seen in the blow-up in the right-hand panel. It is claimed that, whenever oscillations with period three occur, the system can be driven into the chaotic regime by shifting its parameters.

8.3 Turbulence Made Simple

One of the problems discussed in Kolmogorov's student seminar in the mid-1950s was to prove the practical impossibility of weather forecasting. We all suspect this, of course, but we may think that weather is poorly predicted just because atmospheric turbulence is too complex even for the fastest computers to crack. For Edward Lorenz, a meteorologist, the forecast problem was particularly acute, and he undertook to demonstrate that it would fail even in the simplest setting. Lorenz (1963) built up his model by expanding the equations of thermal convection in a Fourier series and picking the amplitudes of just the three lowest modes (at least three are needed for chaos). The system has two symmetric equilibria, which can be visualized as convection cells with a clockwise or counterclockwise flow. As the parameters of the Lorenz system are shifted, these equilibria become unstable and oscillations begin. Later on, the amplitude of the oscillations grows and eventually the trajectories start to switch erratically between circling either the positive or negative unstable equilibrium point, as shown in the left-hand panel of Fig. 8.7. This is when chaos begins.

Here again, it is advantageous to construct a Poincaré map, which can be used to follow the behavior in the oscillatory and chaotic regime. As in Rössler's system above, this map, shown in the right-hand panel of Fig. 8.7, is one-dimensional, but it looks quite different, as it has discontinuities corresponding to jumps between

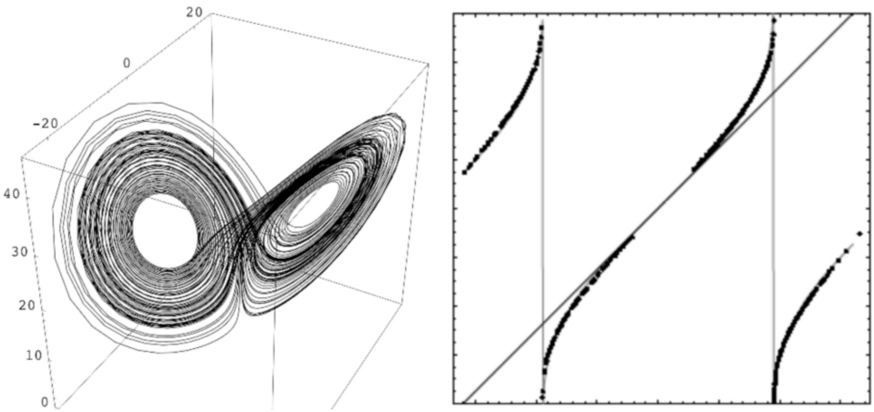


Fig. 8.7 *Left*: Trajectories of the Lorenz system. *Right*: Poincaré map

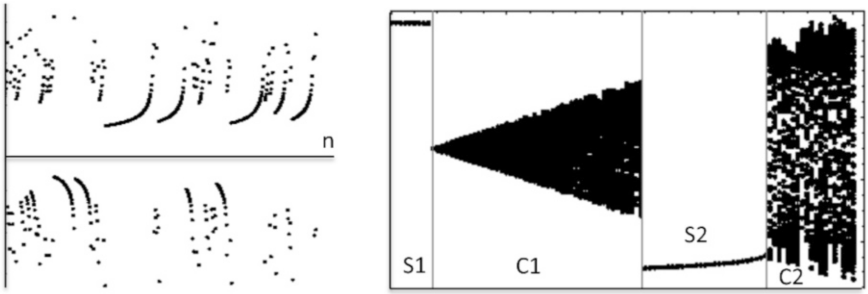


Fig. 8.8 *Left*: Iterations of the corresponding Poincaré map. *Right*: Successive transitions from order to chaos

circling one equilibrium point or the other. This map can be approximated by a relatively simple function (though not as simple as any featured in Sect. 8.2) and iterated many times to get a sequence of points, as in the left-hand panel of Fig. 8.8. Transitions in behavior can be followed by changing a parameter of this approximate map as shown in the right-hand panel. In the interval S1, all trajectories converge to the nearest fixed point of the map, which corresponds to a periodic orbit surrounding an unstable equilibrium. This changes to a chaotic attractor of limited extent in the interval C1. As this attractor spreads out, it suddenly breaks down, and the trajectories starting near either fixed point become attracted to the other one in the interval S2. Later on, there is a transition to a widely extended chaotic attractor, covering the interval C2.

The behavior close to the two boundaries of the interval S2 is particularly interesting because convergence to the stable fixed point occurs after a long chaotic transient, as in the left-hand panel of Fig. 8.9. On the chaotic side of the two boundaries, the trajectory may be trapped near the fixed point, already unstable, for a considerable time before escaping to the chaotic attractor, as in the right-hand panel. This is the Pomeau–Manneville scenario of a transition to chaos through intermittency (Pomeau and Manneville, 1980), quite different from the period-doubling scenario.

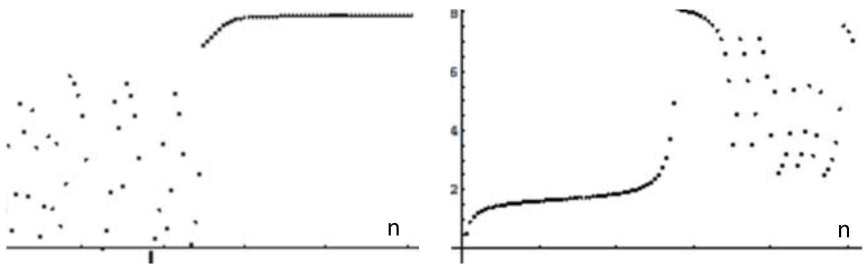


Fig. 8.9 Pomeau–Manneville scenario of transition to chaos through intermittency

Emphasizing the utmost sensitivity of chaotic dynamics, Lorenz coined the term "butterfly effect", using the metaphor of a tornado caused by the flapping wings by a distant butterfly several weeks earlier.

8.4 Complexity in Model Equations

Midway between discrete and continuous toy models and equations describing specific processes stand partial differential equations of a generic kind that have a simple symmetric structure.The first notable generic chaotic model was the Kuramoto–Sivashinsky (KS) equation, derived independently at the two ends of the world and in two different contexts. Kuramoto and Tsuzuki (1976) were working in the context of chemical oscillations and derived this equation to describe perturbations of an unstable front in the FitzHugh–Nagumo system that we met in Sect. 7.2. Sivashinsky (1977) derived the same equation in a more elegant way while studying instabilities of combustion fronts. His paper was submitted to the Journal of Fluid Mechanics in 1975 and was rejected after prolonged reviewing. Incidentally, Feigenbaum's winning paper (Feigenbaum, 1978) was also originally rejected. This says something about the journal reviewing system, which is also chaotic. Everybody complains about it, but nobody knows what can be done, except perhaps replacing scientific journals by web archives.

Continuing with this story, this equation was first published in 1974 in obscure proceedings at Perm University by Alexander Nepomnyashchy, now a professor at Technion in Haifa. He derived it on the basis of general considerations, but did not see that its solutions were chaotic. The chaotic behavior of the KS equation is very mild, lacking the jumps and bifurcations of the models presented earlier in this chapter. As Gregory Sivashinsky once said, watching its simulations lulls you. However, it possesses the main feature of chaos. On the one hand, it is sensitive to initial conditions, and on the other, it forgets them during the course of its evolution. It is no wonder that different routes have brought people to the same equation: it is a general equation describing weak instabilities in propagating fronts.

Perhaps the richest model system is the Complex Ginzburg–Landau (CGL) equation. It has really nothing to do with either Landau or Ginzburg who only worked with its much less interesting *real* version. The equation with a complex variable ("order parameter" in Landau's terms) and complex coefficients appears in the expansion of *any* nonlinear system near the point where it undergoes an oscillatory instability. It generates a whole "world" [or so proclaims the title of a review by Aranson and Kramer (2002)] of instabilities and dynamic patterns, including spirals, their interactions and distortions, and transitions to the various modes of chaotic behavior.

The high point of general theories of patterns and chaos was reached in the mid-1980s to mid-1990s. Common techniques were applied to non-equilibrium structures of different physical origin. Space-dependent amplitude equations and phase dynamics led to an understanding of the genesis and behavior of realistic patterns, constrained by boundaries and blemished by defects. The NATO Special Program on Chaos, Order and Patterns did much to unify the approaches physicists were using and diffuse the resulting knowledge, and it was no accident that its end coincided with the publication of a comprehensive review of the field by Cross and Hohenberg (1993). My own book (Pismen, 2006), written with the intention of collecting together everything I knew before I had time to forget it, came too late, already at the ebb of general theories and model equations. Most of the pictures in Sect. 8.2 and Sect. 8.3 are taken from there.

The turn of the 21st century brought about a more detailed look at complexity. Attention turned to specific applications. Forcing and control of patterns, either enhancing or suppressing the complexity of behavior, were studied in detail. As the humble laptop turned into a supercomputer, more fascinating patterns, the envy of abstract expressionists, were generated by model equations of increasing complexity. The most fashionable direction became the study of *active matter*. Models of active matter can be continuous, similar to the usual equations of continuous mechanics but with added terms expressing internal energy sources; but they can also be discrete, built upon the interactions of mobile particles. The latter models are classified as “dry” when the particles interact directly and “wet” when the interactions are carried by the surrounding medium. An all-round review was published by a group of seven authors (Marchetti et al, 2013).

There is a broad range of applications. At one end of the spectrum are particles obeying common physical laws and driven by external energy inputs. The simplest (though still very complicated) case is motion of grains in a vibrating layer, similar to Faraday’s experiment (Sect. 7.1), but with a sand bath or a batch of small rods replacing a fluid layer. This is an example of a “dry” system. Colloidal particles or droplets may be self-propelled in different ways. The so-called “Janus” particles are two-faced like the eponymous Roman god, although their two faces do not look into the future and the past; they merely have different properties. One of the faces may be catalytic (Valadares et al, 2010), and the reaction on this part of the surface causes an asymmetric concentration distribution in the surrounding solution that propels the Janus particle by the diffusiophoresis effect (Fig. 8.10 left). Liquid droplets changing concentrations in the surrounding fluids can be attracted or repelled due to Marangoni flow caused by emerging surface tension gradients (Golovin et al, 1995), as schematized in the center panel of Fig. 8.10. These systems are necessarily “wet”.

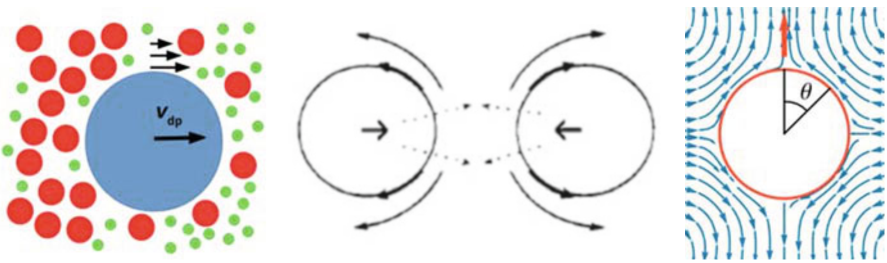


Fig. 8.10 *Left*: Diffusiophoresis. *Center*: Droplet interaction caused by mass transfer and Marangoni flow. *Right*: Stokes flow around a self-propelling particle

Next come natural and artificial microswimmers driven by their own source of energy. Any such swimmer perturbs the fluid in its vicinity, and when it is small and slow, inertia is negligible and the flow is viscous. Its pattern depends on the way the swimmer is propelled. The one in the right-hand panel of Fig. 8.10 is a "pusher", retaining its spherical shape; for a "puller" the direction of the flow is reversed. Different flow patterns appear when the swimmer changes its shape, as bacteria usually do; this is also the means of locomotion of artificial nemato-elastic swimmers (Sect. 5.6) or miniature magnetically driven nanopropellers (Schamel et al, 2014). Flow induced by one swimmer affects all others, so they interact in a very complex way.

Swarms of larger swimming or flying creatures, like insects, fish, or birds, cannot be treated mechanistically, and utterly simple "dry" models have been devised for their description. The simplest model of all (Vicsek et al, 1995) involves only one rule: all "particles" move with the same speed in the average direction of "particles" in their vicinity, being perturbed only by some random noise. One of the authors, the late Eshel Ben-Jacob, who actually came up with the idea, didn't even want to be involved in the publication as the model was so primitive, imitating a kind of magnetic alignment but in a dynamical context. Nevertheless, the Vicsek model successfully predicts a non-equilibrium phase transition from a disordered state at low density or strong noise to an ordered, coherently moving state at high density or low noise strength, and its simulations may look superficially like a flock of birds, as shown in Fig. 8.11. The paper has been cited profusely and has been amended by more realistic details, like short-range repulsion and long-range attraction, which set a certain average density of "particles", so different in the two panels of Fig. 8.11. There is a certain tendency here. The more complicated the system, the more schematic and computationally easy it is the model.

Simple "dry" particle models are converted to continuum equations either by just collecting all the terms allowed by symmetry in the manner of Landau's theory of phase transitions (Sect. 5.2) or in a more quantitative way by using the tools of statistical physics to coarse-grain the model and thus obtain equations describing the dynamics at much longer scales than the particle sizes. Other continuum models have been derived by inserting terms expressing internal activity into typical equations of continuum mechanics (Jülicher et al, 2007).

Fig. 8.11 A flock of birds and a shoal of fish

Devising and studying models is a challenging intellectual game, at least as long as it does not just become a matter of carrying out stereotyped computations for the purposes of producing nice pictures for display. Active matter models are particularly motivated by biological applications, something viewed as central in this century. But how closely do they approach true biological complexity? We shall say more about this in Sect. 9.2.

Chapter 9
Complexity Strikes Back

9.1 Turbulence

I recall seminars where the lecturer started with a slide of a drawing by Leonardo like the one shown in Fig. 9.1, and promptly moved on to the logistic map (Sect. 8.2). Among all enjoyable chaotic models, real hydrodynamic turbulence stands as a mighty fortress.

An apocryphal quotation from Werner Heisenberg (who worked on turbulence for his PhD) goes as follows:

> When I meet God, I am going to ask him two questions: Why relativity? And why turbulence? I really believe He will have an answer for the first.

A similar witticism is attributed to Horace Lamb, author of a classic hydrodynamics textbook:

> I am an old man now, and when I die and go to heaven there are two matters on which I hope for enlightenment. One is quantum electrodynamics, and the other is the turbulent motion of fluids. And about the former I am rather optimistic.

"Enlightenment" supposes a simple answer to a complicated question, like a rabbi challenged to explain the essence of the Torah while standing on one foot. Why should we think that it might be possible at all? No fundamental questions are involved here; no one doubts the Navier–Stokes equation which gives a good description of the hydrodynamics of a Newtonian fluid moving much more slowly than the speed of sound. The problem of turbulence used to have some philosophical undertones, but the studies of simple chaotic models we talked about in the last chapter have already demonstrated that Laplace's demon would have a hard time predicting the future even without the intervention of quantum uncertainty. What remains is the practical need to compute ubiquitous turbulent flows.

Classical fluid mechanics did not touch upon turbulence, considering it to be "hopelessly complicated", in Prandtl's words. The turning point was the understanding that turbulence can be treated statistically. This was first realized by Geoffrey

L. Pismen, *The Swings of Science*,
https://doi.org/10.1007/978-3-319-99777-3_9

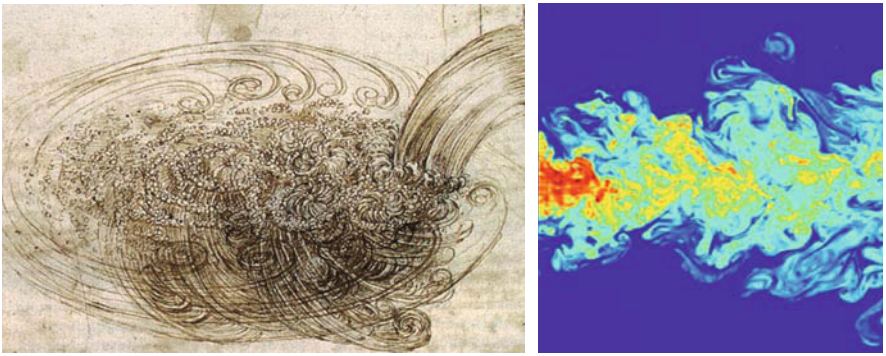


Fig. 9.1 *Left*: Drawing by Leonardo. *Right*: Flow visualization of a turbulent jet

Taylor (1935), whose name is carried by the appellations of the various hydrodynamic phenomena he studied. Restricting to the most transparent case of homogeneous isotropic turbulence, Taylor completely bypassed the traditional approach to turbulent flows: instead of trying to characterize them in traditional straightforward ways, which is indeed hopeless, he suggested trying to measure the correlation functions between velocity components at different times and different points in space.

The breakthrough in statistical theory was the idea of an energy cascade put forward by Andrey Kolmogorov, one of a small number of prominent modern mathematicians inspired by physics rather than pure logic. Turbulence involves motion on many different scales. We see the various features of all sizes, both in Leonardo's drawing and a picture of real flow in Fig. 9.1. The fluid can be stirred on a large scale, and this energy input will trickle down to shorter scales, in the same way that right-wing politicians think that enriching the rich will trickle down to benefit the poor. This is, of course, far better justified in fluid mechanics than in economics. Simply by means of dimensional analysis, Kolmogorov (1941) derived the law of energy distribution among the modes with different wavelengths: a simple power law with the exponent 5/3. By this law, the longer modes are indeed richer in energy, and the energy is steadily transferred to shorter scales until it reaches the limit where viscosity is felt, and is dissipated. Lewis Fry Richardson (1922) expressed this in a paraphrase of Jonathan Swift's sequence of fleas:

> Big whirls have little whirls that feed on their velocity,
> and little whirls have lesser whirls, and so on to viscosity.

The famous physicists did not fair well in the theory of turbulence. In his PhD thesis, the young Werner Heisenberg studied the transition to turbulence through small perturbations of laminar flow. In 1945, while they lingered in detention, he returned to working on turbulence with Carl Friedrich von Weizsäcker, fellow leader of the German nuclear bomb project, apparently without knowing of Kolmogorov's paper, although he did cite it in a later publication [Heisenberg (1948)]. Lars Onsager published the power law with an apparently incorrect exponent in an abstract of a talk in 1945. Lev Landau was tuned to generic phenomena, and might have dis-

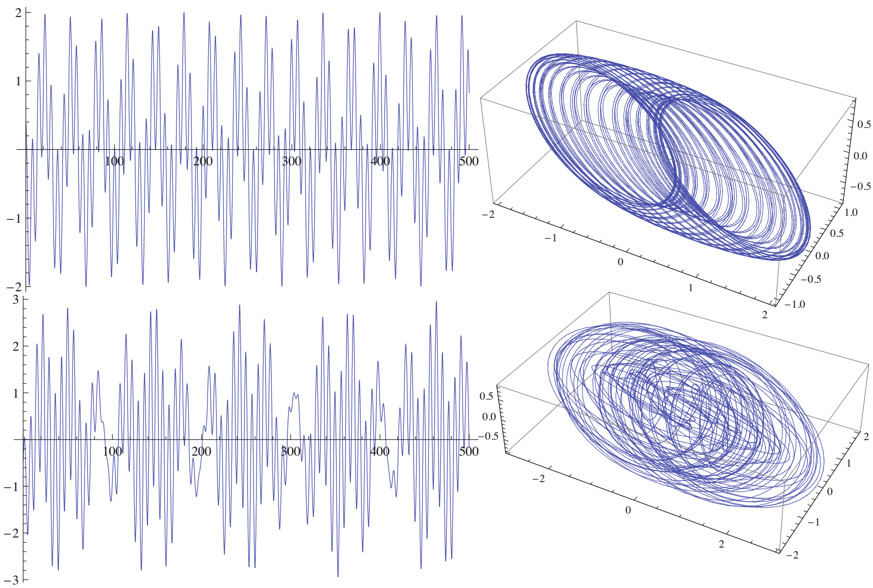


Fig. 9.2 Quasiperiodic motion: superposition of two (*top*) and three (*bottom*) oscillating modes

missed as "sick" some curiosities regarding the behavior of complex systems which now find their way onto the cover of Physical Review Letters and Nature. He envisaged the transition to turbulence as the gradual excitation of a large number of waves. One can see in Fig. 9.2 that a superposition of two oscillating modes already looks complicated – this is just a quasiperiodic motion. Mixing three modes, as in the lower panel of the same figure, may look chaotic to the eye, but it has nothing to do with real chaos.

Kolmogorov reportedly chuckled at Landau's scenario, saying that he was probably unaware of more complex dynamical systems, known already to Poincaré, where chaos arises due to nonlinear effects. We discussed the chaos occurring in toy models in Sect. 8.2: even very simple mechanical systems, like two coupled pendulums, can exhibit unpredictable chaotic behavior. Kolmogorov (1954) asserted that quasiperiodic motion, as in Fig. 9.2, is unstable. This was later rigorously proved by Vladimir Arnold and Jürgen Moser and is known as the KAM theorem.

The ensuing story of the statistical theory of turbulence was difficult and less than brilliant. Kolmogorov's power law is followed quite closely in experimental studies of (almost) isotropic turbulence (which is not so easy to implement). There were a number of attempts to explain small deviations in different ways, all inconclusive. When trying to compute correlations in a nonlinear system, we are driven to correlations of higher and higher order; the same problem is encountered in the BBGKY hierarchy in the theory of liquids (Sect. 5.2). This infinite progression has to be cut off by expressing higher correlations in terms of lower ones. Various "clo-

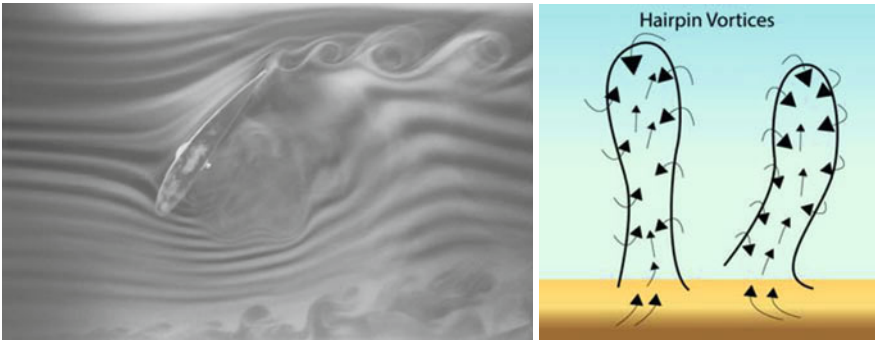


Fig. 9.3 *Left*: Visualization of an airfoil in a wind tunnel. *Right*: Drawing of hairpin vortices in the boundary layer

sures" of this kind have been suggested. These are used in computational models of commercial fluid-mechanical software, but are not always reliable. Diagram techniques similar to the one developed by Richard Feynman in quantum electrodynamics (Sect. 1.5) were applied. I recall one distinguished scholar saying about another one that the latter's merit was in bringing statistical methods to a conclusive dead end. Nevertheless, work in this direction is still pursued.

Practical problems are far from being isotropic and homogeneous, they involve inflows, outflows, and boundaries. Although no reasonable physicist attempts precise prediction, statistical descriptions are satisfactory only as long as they do not concern us personally. The challenging feature of turbulence is that it is not plain random, but has elements of structure, as can be seen in Fig. 9.3 and again in Fig. 9.1. Studies of toy models encouraged an alternative approach to turbulence: discerning *coherent structures* which follow relatively simple dynamic equations and remain more or less persistent while evolving in time. The vortices and vortex streets seen in these and many other pictures are examples of structures of this kind, and more ephemeral patterns can be discerned in familiar cloud shapes and whirls of spreading smoke. Methods similar to face recognition algorithms were developed to distinguish these structures in turbulent flows (Sirovich, 1987). However, nobody has ever given a working definition of coherent structures, and their persistence time is usually quite limited.

Computing flows is still not the end of the road. For many applications, it is important to know how turbulence mixes different chemical species in the fluid (recall Fig. 3.4). This problem is still more complicated, especially in liquids where diffusion is slow and patches transported by turbulent flow have to be dispersed still more finely by diffusion in viscous layers on the way to their destination, which might be a catalytic surface or a rapid reaction front. Further complications are brought about by evaporation and condensation processes and heat transfer, essential both in chemical plants and in the atmosphere. Although we doubt that even God knows turbulence theory, practical computations of turbulent flows are gradually improving. We keep listening to weather forecasts with attention, and forgive failures, taking note that short-range forecasts gain reliability as computers become more powerful and data banks swell.

9.2 Mechanics and the Chemistry of Life

Muscles and bones work in ways that are not unlike those of mechanical devices, provided we do not look at their internal structure, and this was the attitude of Leonardo da Vinci, engineer and artist, when he designed a machine to imitate bird flight. Robotic-looking drawings of animals and people by Giovanni Alfonso Borelli (1680) show the same tendency to reduce motions of the animal body to mechanical principles (Fig. 9.4). The heart has been firmly established as a muscular pump driving the blood circulation network at least since the work of William Harvey (1628).

It was a long way from there to attempts to apply the laws of continuum mechanics to the functioning of living tissues, an enterprise first undertaken in the last quarter of the 20th century. Modern imaging methods have revealed the essential details of cell structure. We are only concerned here with those relevant for mechanics and transport, so let's keep it simple. There is a nucleus containing the genetic material – DNA, the cytoplasm where the protein chemistry unfolds, and the membrane connecting the cell with, and separating it from, the outside world. A separate nucleus is present in *eukaryotic* cells of the kind found in all higher life forms, both animals and plants. These cells have already existed for two billion years and also have various organelles in the cytoplasm which are believed to originate from symbiosis with lower (prokaryotic) life forms. The most important of them are mitochondria, which synthesize the ATP molecules used as an energy supply for the cell. What is important for our narrative is the structural component of the cytoplasm – the cytoskeleton

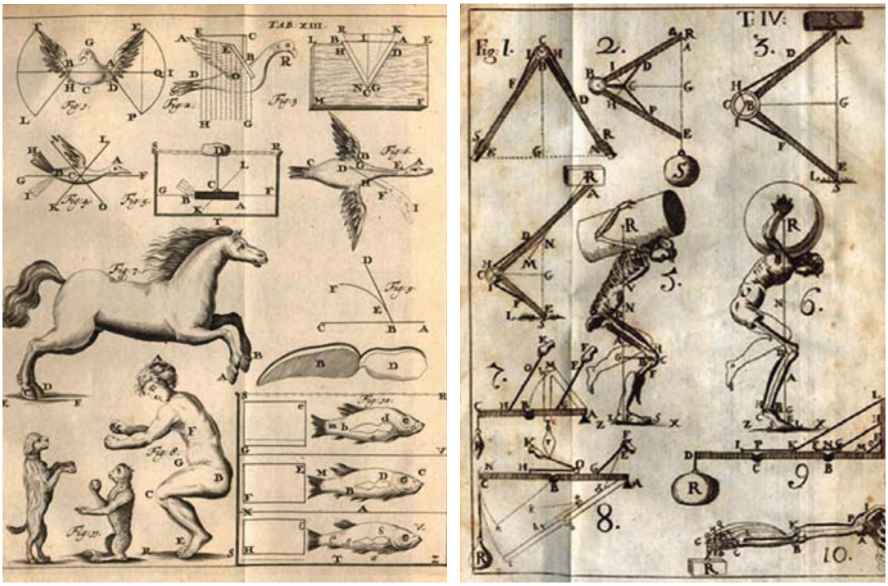


Fig. 9.4 Drawings by Giovanni Alfonso Borelli

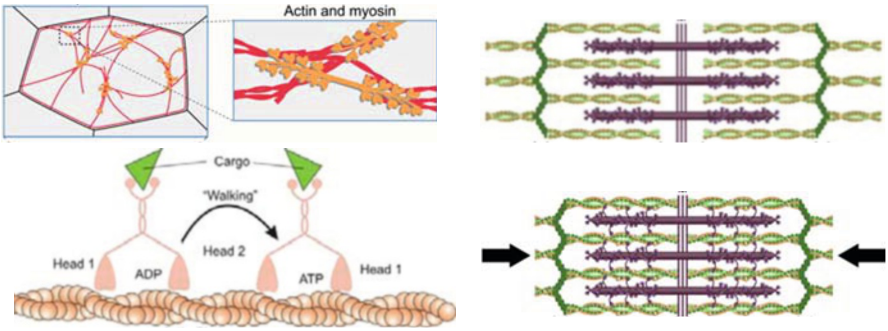


Fig. 9.5 *Top left*: The network of actin filaments within a cell and a blow-up showing their cross-linking by myosin motors. *Bottom left*: myosin motors "walking" on an actin filament carrying their cargo. *Right*: Relaxed and contracted sarcomere structures in the muscle

which, as is already clear from the name, keeps the cell together. The structure of the cytoskeleton is sketched in the upper left-hand panel of Fig. 9.5.

The cytoskeleton may include three kinds of filament, but it is enough for us to talk about only one of them, actin filaments, which are the thinnest, most flexible, and most ubiquitous of all. These filaments are alive, permanently growing, dissolving, and breaking. They are cross-linked and stressed by myosin molecular motors. Besides holding the network together, the molecular motors have another function: they carry cargo – protein molecules, like mail parcels, to their destinations, literally walking upon actin filaments, as they repeatedly change their conformation, lifting and lowering their "legs" (in fact, called "heads"), as shown in the lower left-hand panel. The motors are fed by the energy stored in ATP molecules. Actin filaments and myosin motors, also assembled into filaments, are neatly organized in the *sarcomeric* structure of muscle cells which are contracted due to the motor action, as shown in the right-hand panel of Fig. 9.5.

The cellular cytoskeleton is also essential for the mechanical connection with other cells, to a substrate, or to the extracellular matrix through attached integrin molecules or focal adhesions. The latter are built rather like the automobile clutch, in several protein layers. Unicellular microorganisms move in more variegated ways than the higher animals in Borelli's drawings, and animal cells reshape and move in a way that is not so different from them, most prominently during embryonic development, but also in muscle contraction, wound healing, tumor spreading, and all kinds of tissue rearrangement. Biologists study cell motion *in vitro*, in cell cultures – much easier than *in vivo* – by placing them on a substrate. The most common, but not exclusive way of moving is via polymerizing actin filaments at the leading end and then breaking them behind, with actin monomers transported from one end to the other. Just imagine yourself getting around in this way! It is not fast-moving, micron after micron, but they are not in a hurry: focal adhesions, seen as tiny dashes in the upper right-hand panel of Fig. 9.6 and at a larger scale but without detail below, have to be broken behind and built up ahead. Note that all cellular structures are dis-

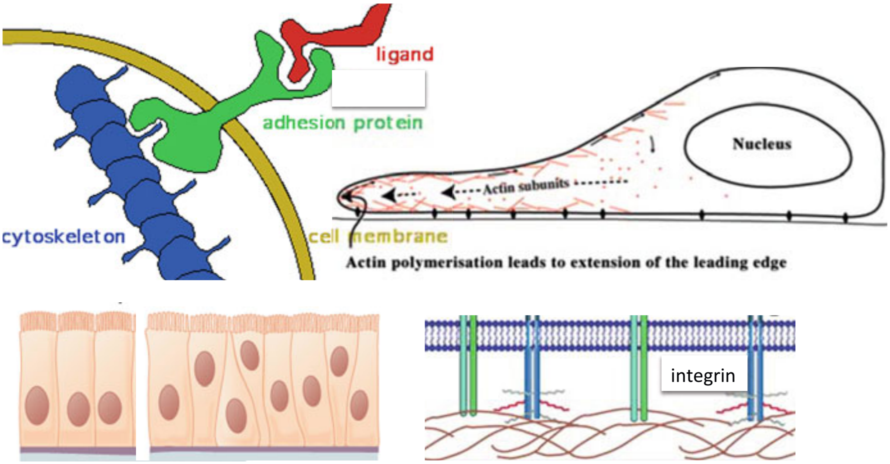


Fig. 9.6 *Top left*: An adhesion site. *Bottom left*: Cells in an epithelial layer. *Top right*: Cell migration by polymerization at the leading edge. *Bottom right*: A cell membrane attached by integrin molecules to a substrate or to the extracellular matrix

pensable; they are dissolved and synthesized as the need arises, while the building material is conserved.

What can the theoretical physicist do about this unfathomable complexity, which we have only lightly touched upon here? There is a lot of activity, including an ever growing number of publications, reviews, and workshops, but there have been no breakthroughs to speak of, only a sense of a gradually growing understanding. Polymer scientists may feel at home with actin networks: like rubber (Sect. 5.6), they are cross-linked polymer networks. The differences are huge, however. Actin filaments are stiff. Their "persistence length", that is, the distance over which the orientation of the filaments changes considerably, is much greater than the distance between crosslinks, so they are almost straight in-between. This actually makes it easier to estimate their elasticity and viscosity, but only if we can neglect interactions between different segments. The result comes out strongly nonlinear: the filaments strengthen as they are stretched – till they break. On the other hand, actin filaments cannot sustain compression, and more sturdy structural elements – microtubules or the same actin filaments bundled up into stress fibers – have to do this job.

Complications come in when we recall that, unlike rubber, the cytoskeleton is *alive*. The myosin motors keep attaching and detaching, creating a random flickering that is much stronger than mere thermal noise – it is as if, as far as fluctuations are concerned, the cells are kept at a much higher temperature than a thermometer would measure. Moreover, this noise is not as random as thermal noise; as in hydrodynamic turbulence, the energy is not evenly distributed between different frequencies. The flickering can be modeled relatively easily when the filaments are organized in a regular structure, as in muscles, giving pictures like the one in Fig. 9.7. It is due to this ceaseless activity that it is hard, say, to keep an outstretched hand

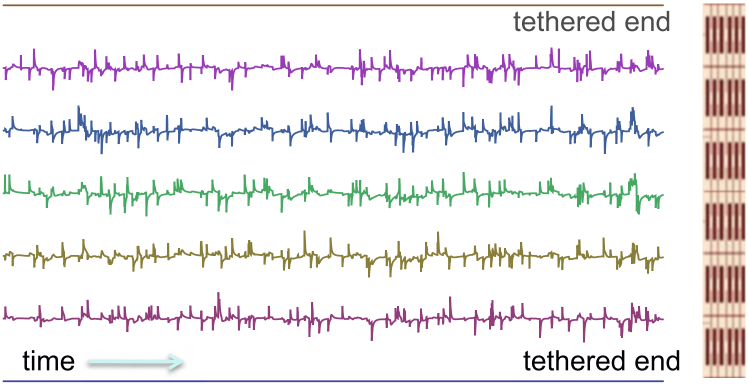


Fig. 9.7 Fluctuations in a muscle at rest, modeled as a sarcomeric structure shown *on the right*

without doing any work: the motors consume energy even when they just wander around. In a disordered network, this flickering causes random rearrangements; if the network is detached from the cell membrane, as is often done in *in vitro* studies of actin–myosin mixtures, it collapses.

Another complication is permanent growth and decay. This is a necessary feature of a living structure; as noted above, cells move in this way. The membrane of animal cells does not support actin filaments mechanically. It is too feeble for that. Instead, this is where monomer units are attached to the filaments to be carried, as on a conveyor belt, to the filament's end, where they are released to diffuse and get attached again. The attachment, detachment, branching, cross-linking, and occasionally break-up of filaments depend on many factors, both mechanical and chemical. This is good for the cell, for the flexibility of its multiple tasks, and this is why evolution has created this sophisticated mechanism, which is such a headache as a modeling problem. There are nice simple models elucidating parts of the big picture, in particular, the way mechanical stress affects polymerization at the membrane end, something which affects the motion of the cell and its communication with neighbors (Mogilner and Oster, 2003). However, complex chemical factors remain out of reach of such mechanical models.

Setting the details aside, the physicist may "coarse-grain" the problem, as is commonly done in all applications, and view the cytoskeleton as a viscoelastic continuum characterized by certain parameters, estimated theoretically or measured. Then one can make a little headway, especially when treating not a single cell but a *tissue*, like an epithelial layer consisting of many adhering cells. There is a risk of losing special features of living matter in this way, but this can be avoided by inserting some terms accounting for "activity", as in the active matter models mentioned in Sect. 8.4. The necessary element of these models is the *polarization* of cells. Both their internal elements and cells themselves viewed as a single unit, either isolated or as part of a tissue, have a certain orientation determining their force field and/or direction of motion. The physicist is most at home when the cell can be treated as an elementary object, or a kind of "atom". There are situations when this is justified. Thus, cells on an elastic substrate can just be treated as force dipoles. In this way, it is possible to explain the interaction between cells through the substrate, which they deform, and their migration to stiffer areas of the substrate, which is called durotaxis (Safran et al, 2005).

Still more radical means of simplification are being put to the test. With the decline of theoretical nonlinear science studying generic models (leaving it to mathematicians to pick up the crumbs), many refugee physicists have moved to biophysical problems. It is rather easy to use a modification of a generic model to produce a moving active spot and interpret it as a moving cell; models of this kind have been published more than once. One resourceful physicist has suggested a "modular" approach: you start from a simple basic model and add details as the need to better imitate reality arises. This reminds me of a tale about a soldier promising a peasant woman to cook a soup out of an axe. Then, as time goes on, he asks the woman to add more and more nutritious ingredients, so that in the end the soup comes out tasty! The problem is that it might be possible to imitate some features of a complex system in a way that has nothing to do with the actual mechanism.

The ultimate complications are brought about by *chemistry* or, in biological terms, the interactions of the various enzymes and other species given arbitrary names by biologists. The chemistry of life is immensely complex, maybe unnecessarily so – but Nature is not aware of Occam's razor, and cares for robustness and stability more than for rational and economic design. The left-hand panel of Fig. 9.8 shows the metabolic network of a particular process used by a particular plant to release stored energy. The lines show the interactions between enzymes and metabolites marked by dots. Actually, citation networks and many other networks look something like this, but otherwise there is far less universality here than the physicist would like to see.

Commonly, whenever chemistry and transport of proteins are essential, no more than a single species is taken into consideration in a mechanical model. There is usually no point in trying to include more since, even when detailed mechanisms like the one above have been guessed, the reaction rates are unknown, and biologists have neither the tools nor the intention to measure them. However, combining mechanics with chemistry is not futile. A good example is the treatment of facilitated transport through the cell membrane (right-hand panel of Fig. 9.8). Membranes are highly specialized: their task is to let in what the cell needs and let out what it has to get rid of, whilst all the time guarding against invaders. The opening and closing of active channels is largely regulated biochemically, but mechanics may play a role, causing channels to open or contract when the membrane is stretched or compressed, and a feedback loop may even arise when the transmitted ions or other agents affect the deformation of the membrane (Salbreux et al, 2007).

Some researchers concentrate on particular details, say, on the mechanics of protrusions at the leading edge when the cell moves. There are also attempts at very detailed modeling, like the motion of every actin monomer and every myosin motor in the network – but even this is still a long way from the molecular detail that may sometimes be decisive. We are coming here to the general problem of reduc-

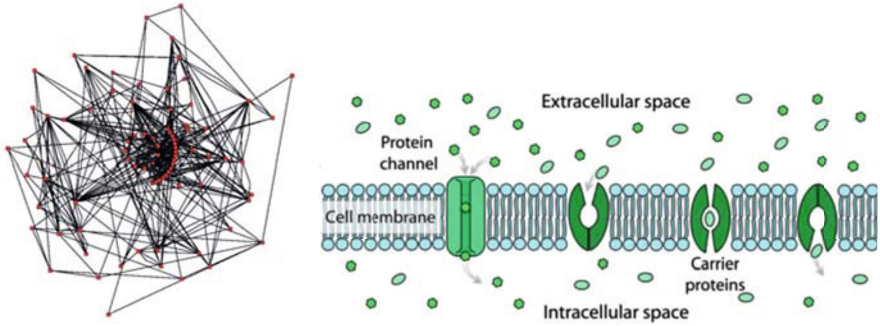


Fig. 9.8 *Left*: A metabolic network. *Right*: Facilitated diffusion through a cell membrane, showing an ion channel *on the left* and three carrier proteins *on the right*

tionism versus emergence. Ideally, one would like to be able to show how emergent properties arise from the interactions of lower-order entities. This is what statistical physics manages to do, but only for media with relatively weak or well-ordered interactions. In biophysics, it may work only with great ingenuity and good luck, and in a fairly limited way. On the other hand, a holistic view of the organism brings nothing but highfaluting words, so the only way is to struggle with minute details and hope for illumination to come our way, even if it lights up only a tiny spot.

Much research here, as in other fields of science, is driven by curiosity and/or the inertia of academic life under publication pressure, but the potential benefits for human well-being might be greater than elsewhere. When my colleague, reporting the results of our funded project, persuasively presented a model of a spreading cancerous tumor, a woman in the audience asked whether a medical practitioner would now have to learn mathematical modeling. Doctors would certainly get help if it comes to it, but we are still a long way from realistic models, and it is not clear whether we will ever come to this level of detail. Biologists also have a long way to go. Optimists say that some humans already born (presumably, rich) will be immortal, and not in the manner of Jonathan Swift's struldbrugs, but eternally young as well. Researchers can be encouraged by Robert Sapolsky's forecast (Sapolsky, 2017): "immortality is unattainable, so research can be funded forever".

9.3 Heredity and Evolution

One of the ancient questions is the way complex living forms emerge. Beliefs in spontaneous generation of lower forms from non-living matter, e.g., insects from putrefying earth, go back to pre-Aristotelian Greek philosophers. They persisted through the ages in spite of being objectionable from the Biblical point of view, and were subject to inconclusive experimentations when it became fashionable, until finally disproved by Louis Pasteur (1864). Living things develop following a predetermined plan – long believed to be set by the Almighty – but where is this plan hidden? Quite an absurd answer was given by the spermists of the early modern age: they thought there was a tiny homunculus hidden in the sperm (Fig. 9.9). This would engender infinite complexity: would there be a chain of homunculi going "all the way down"?

A far simpler plan was discerned in Gregor Mendel's study of variation in peas, as he laid the foundations of modern genetics while working alone in a quiet monastery garden. His study, published in 1866, established the rules of heredity by combination of dominant and recessive traits which, as it was later understood, are carried by the genes in sperm and eggs. His work remained unnoticed until rediscovered and verified in 1900; he did not live to see the spread of his fame. Gregor Mendel essentially quantized heredity – although there was no connection with quantum mechanics and he did it long before Planck. He proved that inherited traits do not mix: a white-flowered and a red-flowered plant, when cross-fertilized, do not produce pink-flowered progeny. What is inherited is the dominant trait, while the recessive one can reappear with probability 1/4 in an offspring of its two carriers. Recessive genes often carry harmful mutations; this is why custom and religions prevented incest in most cultures for millennia before Mendel, with a few exceptions, as in the Egyptian ruling dynasties.

It was natural to conjecture that there should be a material carrier of inherited traits. It was established by the turn of the 20th century that this was located in the

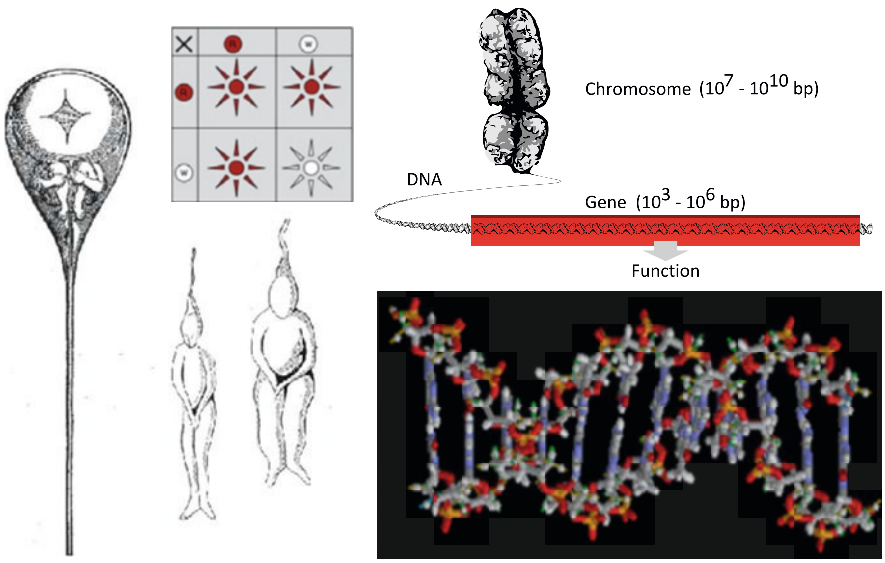


Fig. 9.9 *Left*: A homunculus in the sperm, drawn in 1695 by Nicolaas Hartsoecker. *Inset*: Mendelian inheritance. *Top right*: Genes in a chromosome: bp stands for base pair. *Bottom right*: Double spiral of DNA bridged by the coding base pairs

chromosomes[1] within the nuclei of eukaryotic cells. The carriers were first thought to be proteins, but then desoxyribonucleic acid – DNA – was finally established as the information carrier by Avery et al (1944). Erwin Schrödinger (1945), in his book pioneering physicists' involvement in the problems of biology, justifies at length why the information carrier should be a *molecule*. He followed the work of the geneticists Max Delbrück, later a Nobel Prize winner, and Nikolai Timoféeff-Ressovsky. For the latter, and also the latter's teacher Nikolai Koltsov, this should have been evident, but it was less clear to a physicist accustomed to the idea that macroscopic phenomena should involve a very large number of atoms, since otherwise their ordered structure would be destroyed by random fluctuations. It was already established at the time that genetic traits should be coded by no more than hundreds to thousands of atoms, a minuscule number. In order to keep such a structure intact, strong bonding is needed, and covalent chemical bonds, already understood in the framework of quantum mechanics (Sect. 6.4), were the only way to keep the frequency of mutations low. Schrödinger called DNA an "aperiodic crystal", much more interesting than the periodic crystals commonly studied by physicists. Indeed, only an aperiodic structure can carry information.

To put this in perspective, we need to keep in mind the technology of memory storage at that not so remote time, within the range of a human life. The most compact man-made memory device was still a *book*. Try to count the number of atoms in a single printed word – it will be hard. The memory elements of the first computers were perforated cards, and ropes were used in the rockets shuttling to the Moon in the late 1960s. Even the far more compact modern memory elements are still macroscopic, and further miniaturization is impeded by fluctuations, exactly as Schrödinger had feared. Future technology may imitate Nature. Goldman et al (2013) demonstrated the use of synthetic DNA for information storage. It may become practical only if efficient ways of writing and reading this code are found. Nature knows them, but they are not as fast as those we are accustomed to in modern electronic devices. Nobody would want chemical synthesis to run in their smartphone in the same way it runs in their body. The closest we are now getting to the way Nature operates is 3D-printing directed by a stored program – but the material used is produced elsewhere. Another difference is that we are in much more of a hurry than Nature is: it operates on quite different time scales.

The structure of the genetic code was finally established by Watson and Crick (1953), rushed by a perceived competition with Linus Pauling. The "homunculi" turned out to be tightly compressed in the double spirals of DNA carrying genetic information. This was the high point of simplicity, the culmination of many long centuries of search and confusion, the triumph of a simple and unique principle beneath the infinite variety of forms. Only four "letters" – bases attached to the spiral backbone – suffice to encode the order of amino acids assembled into protein molecules. The design ensures the stability of hereditary forms, and at the same time is sufficiently flexible to allow new forms to develop by recombination and mutation of genes, as shown in the symbolic artist's rendering in Fig. 9.10. Mutations are

[1] This is an accidental appellation, related to the fact that they are easily stained by some dyes. Many biological terms are as arbitrary, but then just think of colored and flavored quarks.

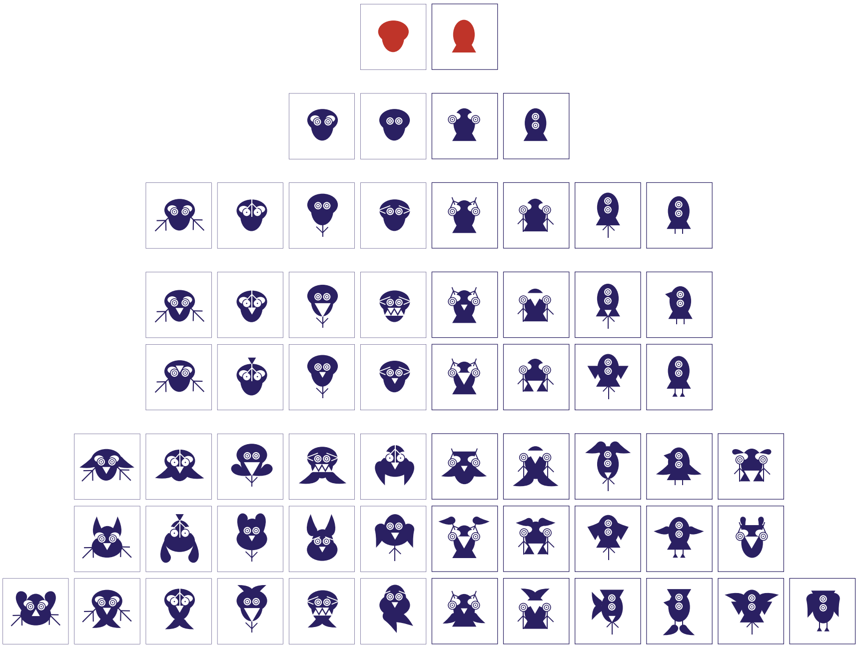


Fig. 9.10 Artist's conception of the evolution of variegated complex forms (Timerman and Timerman, 2002)

rare because chemical bonds can be broken only by strong external inputs, such as radiation or the impact of highly reactive chemicals. The genetic code is carried by all cells of the organism, not only by gametes carrying genetic information to further generations. Their mutations lead to cancer – another reason to keep their occurrence as sparse as possible.

The genetic code is a book written in a certain language which has to be understood. The meaning of this language lies in the proteins it generates. It is not easier to decipher a DNA string than an ancient inscription left by an extinct civilization. As related by Francis Crick (1998), the great physicist George Gamow tried to deduce exactly how this code is structured. It follows from simple combinatorics that, if each amino acid to be inserted in a protein is coded by a combination of overlapping triplets of bases in any order (as sketched in Fig. 9.11), four "letters" are just enough to code twenty amino acids in a unique way. This would immediately explain why there are four bases and twenty amino acids, not more and not less. Alas, Nature is not as rational as the physicist sees her to be! There is no overlap, the order of bases is important, and the code is highly degenerate: some combinations of bases do not code any amino acids, and the correspondence between base triplets and amino acids is not one-to-one. Gamow could not follow Dirac in relying on the esthetic appeal of the theory. He could not, like Einstein, pity God for not following his beautiful design. “Beauty is truth, truth beauty,” said John Keats much earlier – alas, not always.

A close analogy can be found in the Hebrew script where most roots are triplets of letters; their order counts, of course (Fig. 9.11 inset). Some triplets have more than one meaning and many others have no meaning at all. The genetic code is now read with no less ease than the Hebrew Bible, and is computerized for cheap and easy sequencing; the entire human genome is known, and its variations are used for all purposes from following ancient migrations to checking genetic defects of an embryo in the womb to catching criminals. It turns out that the overwhelming majority of the genome consists of “senseless” noncoding sequences, what used to be called “junk DNA”. Although Nature is not expected to be rational, this looks suspicious, and for good reason. The cellular chemical factory operates in a sequence of consecutive stages. Proteins are not synthesized on the DNA; its base sequences are copied on RNA molecules that are sent out of the nucleus into the cytoplasm where the main operations room is set out. And it gradually transpires that noncoding DNA and RNA play an important role in supervising the activities on the production lines of this factory. Nature does not accumulate junk without purpose.

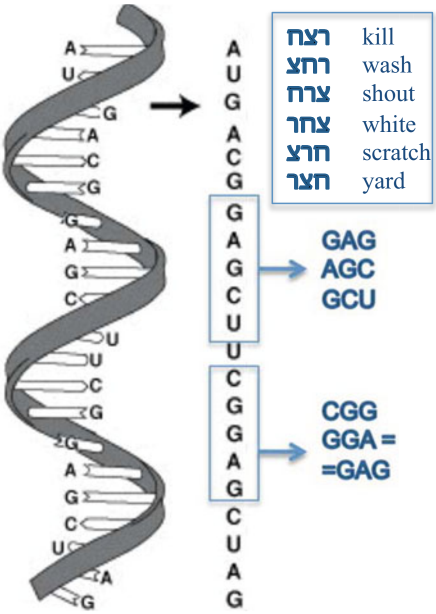


Fig. 9.11 The genetic code as read by Gamow. *Inset*: Permutations of Hebrew roots

Mendelian laws of inheritance also turned out to be not so clear-cut, as was shown by numerous studies of a “model organism” – the fruit fly *Drosophila melanogaster* – initiated by Tomas Hunt Morgan in 1908. If bases are letters and the amino acids they code for are words, the proteins to be synthesized are sentences – but what are genes? They are commonly viewed as segments of the DNA chain (Fig. 9.9), paragraphs of the genetic book, but they are observed as traits, and traits may be determined by several genes. The relevant genes may be separated due to *crossing over*, the exchange of segments of the DNA chain between chromosomes (Fig. 9.12), as envisaged by Morgan.

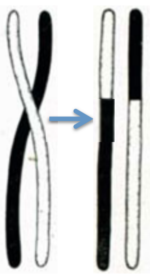


Fig. 9.12 Crossing over

As the clarity of the genetic code gets blurred, so does the mechanism of heredity and evolution. In Sect. 6.4, I mentioned the suppression of scientific genetics in the USSR under Stalin. The opposing force, besides malice and ignorance, was Lamarck's theory of the inheritance of acquired characteristics. Jean-Baptiste Lamarck was a noble and courageous, though erring man, and his teleological idea was noble, too: *le pouvoir de la vie*, striving for organization and perfection, far more beautiful than Darwin's random genetic wandering in the dark, with cruel extinction of the unfit. "Who is the swordsman fighting for the honor of Nature? But, of course, the fiery Lamarck!" wrote Osip Mandelstam. The only problem with this is that it is wrong. There is no way muscles you pump up in a gym or the yield of milk of Lysenko's cows fed by chocolate could ever be passed on to the respective progeny. Subtle exceptions, however, are starting to turn up in *epigenetics*. Some phenotypic (observable, as opposed to innate) traits are due to operational details of the cells' chemical plants. They may feed back on the chemistry of chromosomes and be inherited without affecting the genetic code itself, yet possibly changing the expression of genes by turning them on or off.

Darwin's laws are cruel: it is not the cleverest, strongest, or most beautiful that survive, but the fittest, and the definition is circular: those who survive are proven to be fit. Cockroaches are older than us by three hundred million years, they have survived through mass extinctions and ice ages, and will survive if we extinguish ourselves in a nuclear war. Cancer cells are fitter than benign ones, as they never stop dividing until they die themselves after destroying the cellular society that supports them.

9.4 Development

The generalizing attitude of 20th century physicists has given way in the new century to focusing on details. Complexity keeps creeping in, as thousands and thousands of biologists, biochemists, and biophysicists, driven by the desire to understand the inner workings of life – maybe to be able to extend it eternally – and sustained by grants more generous than in other branches of science (including even the megalomanic edifice at CERN) – have been struggling for more than half a century since the discovery of the double spiral to uncover the detailed mechanisms involved in translating the molecular code into living forms.

The DNA code does nothing more than coding the synthesis of proteins, and it is the latter that do the rest of the hard work. "Few processes are more elegant, I think, than the construction, following the program of DNA, of a baby from a fertilized egg", observed Frank Wilczek (2012). He did not perceive detail while looking down from his ivory tower: the process might be more complicated than creating the Universe in a Big Bang. Freeman Dyson's insight that "scientifically speaking, a butterfly is at least as mysterious as a superstring" (Dyson, 1988) may be rather an understatement. The difference between string theorists and biophysicists is that the latter's fantasies may not even live long enough to be published.

Development does indeed start with a sequence of divisions creating identical cells, just as Tegmark (2014) imagined the Universe to start. But then cells have to be specialized in much more varied and subtle ways than elementary particles, or galaxies, or stars, and the major problem of development is to work out how the inscribed plan is implemented. It should indeed be inscribed in the DNA code but in a very indirect way, as is realized through the exceedingly complex interaction networks of proteins coded by different genes, which remain undeciphered even in the most comprehensive studies. The most obvious riddle is how cells placed at different locations know they should switch on certain groups of genes to implement the functions they are destined to carry out, to become a part of the heart, or the gut, or a finger? In Sect. 7.2, I mentioned the naive idea of spontaneous symmetry breaking by Alan Turing (1952), which may be applicable to chemical patterns in a lab but not to embryos. This should be clear: the observed chemical patterns, as well as patterns in model computations, are repetitive, and whatever variety they have is caused either by a difference in initial conditions or by random inputs. On the contrary, the morphogenetic process is unique and precise, with variegated features emerging at precise locations. Even segmentation, common in animals, does not work by Turing: particular chemical interactions are implemented to build up each segment – which may be the only way to ensure that the number of segments is independent of the size of the animal.

A rival morphogenetic scenario, supported by evidence accumulated over half a century, as described, for example, in the book by the 1995 Nobel Prize laureate Christiane Nusslein-Volhard (2008), has been put forward by Lewis Wolpert (1969). Unlike Turing patterns emerging on a homogeneous background, patterns of biological development are governed by *morphogens* emanating from a certain source, thereby breaking the symmetry of a featureless background, and the positional information is provided by morphogenetic gradients. Repetitive patterns do occur; for example, the formation of fingers has been reproduced by a rather realistic two-dimensional simulation. However, fingers have to be generated at a suitable location, which requires positional information, and in a predetermined number, which requires scale invariance. Moreover, all fingers are different, and this should depend on positional information along their sequence. Even in development processes generating repetitive regular patterns, such us hair follicle or feather formation, patterns do not emerge by being triggered by random fluctuations on a homogeneous background, but are generated by a morphogenetic wave propagating in a predetermined direction.

The Turing and Wolpert scenarios do have a common feature. Activation and repression agents must be combined in both scenarios, and the essential feature of acting genetic schemes is a *feed-forward* motif (Alon, 2007). The simplest patterning scheme involves a single incoherent feed-forward loop $S \to P$, $S \to T$, $P \dashv T$ that includes two activating ($\to$) links with different thresholds, initiated by the same signal S (induced by a morphogen), and an inhibiting ($\dashv$) link from the intermediate protein P to the target T. This scheme generates Wolpert's "French flag" pattern with the target expressed in the central ("white") interval, where the signal level is below a higher threshold of the link to the protein P and above a lower threshold of the direct link to the target. Differences in diffusivities of morphogens also play a role in localizing activation or repression thresholds, although there is no general reason for the latter to be less diffusive. And, of course, all morphogenetic patterns

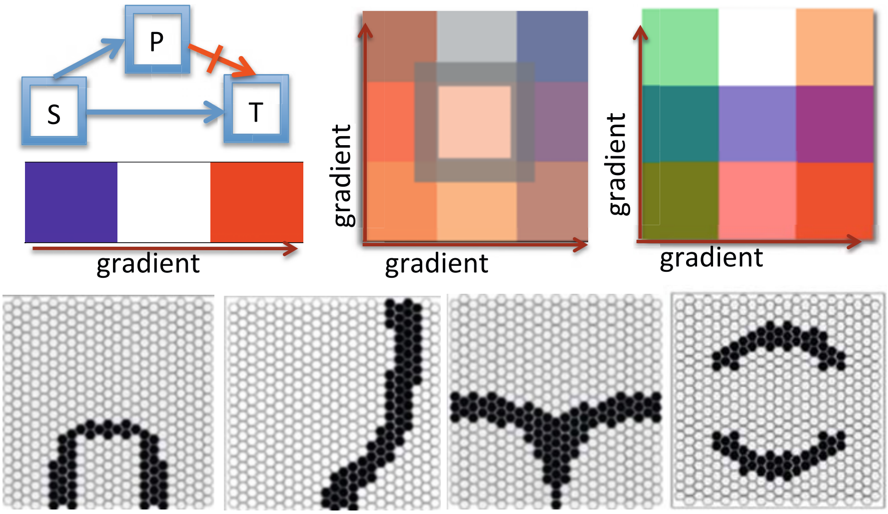


Fig. 9.13 *Top left*: A feed-forward loop and Wolpert's "French flag". *Top right*: Nine expression domains of the target and the ligand. *Bottom*: Examples of patterns generated by modeling different genetic schemes

are dissipative structures in a wide sense, as they are actively driven and sustained far from equilibrium.

The "French flag" scheme can be straightforwardly extended to two-dimensional (2D) patterns under the combined action of crossed gradients. A common example of 2D signaling is found in combined anterior–posterior and dorsal–ventral gradients in a developing *Drosophila* eggshell (Berg, 2008). A second signal, generated by a morphogen with the gradient in the direction normal to that of the first signal, may induce 2D patterning that can be presented, extending Wolpert's simile, as a nine-color superposition of the French and German flags (Fig. 9.13). However, this is far from sufficient to explain the rich variety of locations and shapes of expression domains. Notwithstanding the complexity of intracellular interaction schemes, the variety of persistent expression patterns cannot exceed the limit set by intersections of two-level sets of the two signals. Modifying the form and location of the signal source, e.g., replacing a linear source by a point source, would only change the shape, but not the topology of the expression domains. Adding more initiating links with different thresholds may increase the number of subdivisions, but domain shapes will always be set by the signal level sets, making it difficult to explain less regular gene expression patterns.

A possible way to create variegated shapes of expression domains might be to combine external signals with *autocrine* morphogenetic signaling, initiated within the embryonic tissue by *ligands* whose expression is in turn divided by the local morphogen levels into a nine-color map. If there is a single target gene and a single ligand, there are altogether 16 combinations of target and ligand expression in the presence or absence of the autocrine signal in each domain. The diffusional range of the autocrine ligand is typically shorter than that of externally supplied morphogens. If, for example, a particular genetic network is built in such a way that, in one of the nine domains say, A, of the nine-color “flag”, the target is not expressed but the ligand is expressed, while in a neighboring domain B, the ligand is not expressed and the target is expressed only in the presence of the autocrine signal, the target expression will be observed only in a narrow strip near the boundary where the level of the autocrine signal is sufficiently strong. Conversely, if both the target and the ligand are expressed in the domain A, while the ligand is not expressed in the domain B and the target is expressed only at a low level of the autocrine signal, the target will be expressed everywhere except in a narrow strip in the domain B near the boundary with the domain A. This helps to explain gene expression in domains with convoluted forms. Examples are shown in the lower row of Fig. 9.13 (Pismen and Simakov, 2011). The number of combinations grows exponentially as 2 to the power of $2(n+1)$ with the number n of autocrine ligands, leading to a great variety of expression domains for the same intrinsic genetic scheme.

The scaling and robustness problem is an Achilles heel of both Turing patterns and Wolpert’s flags, and it is not helped at all by adding more signaling species. In these cases, the scale of a pattern is fixed by diffusional lengths of reactants or morphogens, respectively, while in reality the development patterns are scaled by the size of an organism. As an extreme example, a mouse and a giraffe have the same number of vertebrae. A variety of mechanisms have been suggested to rectify this contradiction. Some naive attempts suggested doubling a signal by either a counter-propagating signal or a sink at the opposite edge – an arrangement that becomes forbiddingly clumsy in the case of 2D patterning and in the presence of several morphogens, and is never supported by evidence. The scaling problem can be solved by some kind of global control (Ben-Zvi and Barkai, 2010), which is also known to stabilize localized structures in model reaction–diffusion systems. For example, making the morphogen degradation dependent on some chemical species present in a fixed amount and uniformly distributed in a developing embryo would automatically make the morphogen gradients scale-invariant; it could also act in the same way on all morphogens diffusing in different directions. However, global agents require a fast mechanism for sustaining their uniform concentration, and this cannot be achieved by molecular transport. Long-range mechanical interactions may, in principle, serve to provide global feedback, but the actual mechanism is likely to be far more complicated. Efforts to understand morphogenetic problems in the diffusional framework may eventually prove to be futile. Diffusion is too slow to ensure observed characteristic times for establishment of morphogenetic gradients, and morphogens may be delivered in a more sophisticated way through cell extensions specialized in their transport (Kerszberg and Wolpert, 2007). This implies more intricate and specialized interactions that determine positional information.

Gene expression is followed by spatial transformations creating living forms. Some deformations of cellular layers, like invagination, which create inner tissues at an early stage of embryonic development (Fig. 9.14), can be modeled in a visually persuasive way. Mechanical regulation may complement chemistry in the morphogenesis as well as in the functioning of living cells and tissues. The prospect of applying the general equations of continuum mechanics is certainly appealing to physicists. However, chemo-mechanical interactions involve complex specific mechanisms, and the applicability of continuum theories is impeded by the crowded and irregular microstructure of living cells.

The detailed view tends to be infinitely complex. It is necessary to establish what we really need to know. Some features of the morphogenetic process are amazingly well conserved, from *Drosophila* to higher animals, but curiosity and inertia may also drive researchers into thorough studies with irrelevant results. Details are essential when studying human development and physiology, even at the price of sacrificing our mammalian relatives to save human lives, but not every detailed study of the proteins involved in the formation mechanism of a particular feature of the fly anatomy would contribute to either practical or existential knowledge. Detailed studies help to elucidate general principles as well; in particular, only in this way can we understand the relative roles of diffusion and active transport in shaping morphogenetic patterns. More insight is promised by *in vitro* studies in controlled artificially engineered environments. *In silico* studies are also actively pursued, but their results should be viewed with caution, inasmuch as mathematical models of very complex systems are apt to produce results superficially similar to observations, even when their foundations are far from reality.

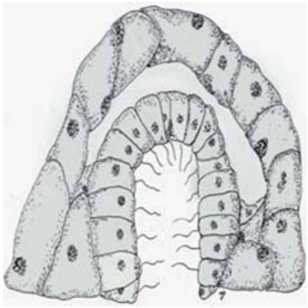


Fig. 9.14 Invagination

Richard Feynman once wrote: “What I cannot create, I do not understand.” Sometimes we create without understanding much, using in-built natural machinery. Selective breeding was practiced for millennia, reducing the wolf to a dachshund among other things, and direct merging of different DNA chains began in the early 1970s. A chance discovery by the team of Jozef Schell and Marc van Montagu (Zaenen et al, 1974) that the *Agrobacterium tumefaciens* microbe induces cancer in plants by inserting detrimental code sequences into the cell’s genome was redirected to beneficial use, adding desired genes, say, to make plants disease-resistant in a cheap and easy way. This led to the development of an entire industry producing genetically modified organisms (GMO), something referred to by its detractors as “Frankenfood”, forbidden in some countries, though its ill effects are not evident. Further advance, more dangerous but still blind, is the synthesis of unnatural DNA containing “letters” different from the four present in the natural genome, initiated in 2002 (Chen et al, 2016).

Real Frankenstein monsters should be truly artificial, perhaps literally made *in silico*, just as the Golem was made of clay. This will very likely be made possible only when manufacturing adopts biomimetic principles of pattern formation governed by signaling, replacing or complementing the common assembly methods (which also encompass, alongside traditional manufacturing, such modern processes as lithography or 3D printing). Intelligent design by humans or artificial intelligence may arrive at simpler and more rational (if not superior) solutions than a blind Darwinian search, and is likely to rely on electric rather than chemical signaling. We do not expect drones ever to be manufactured the way insects are, any more than airplanes are made in the same way as birds or computers as brains, notwithstanding the wonders of natural design – but when the sophistication of intelligently designed autonomous systems reaches the level of their living counterparts, their production process and behavior may also become too complex to be reduced to a few elegant general principles. Will it ever amount to artificial life? Here is an exciting and dangerous prospect: generating a new kind of life or introducing it from elsewhere could be a disaster, like nucleating an alternative vacuum.

These questions are most highly charged when applied to human heredity, development, and evolution. Are we the highest point of creation, never to be changed? This seems doubtful, in view of our imperfections, but we have reached a stage where natural selection is no longer working. Advances in medicine allow people bearing genetic defects to survive and even procreate. On the other hand, it is already possible to select embryos genetically during artificial insemination, which might be a first step toward a society in a dystopian movie where natural conception is forbidden. Genetic modification in the manner of the "Frankenfood" movement to produce "designer babies" is the next step which many will find appalling. Eugenics became a dirty word following the outrageous excesses of the Nazis. But this is still not the end; there are prospects of cyborgs combining biological and electronic components. Would this be the way to *Homo deus* replacing us (Harari, 2017)? To quote De Duve and Patterson (2010): "Whatever the cause of the extinction of our predecessors, it most likely was associated with some kind of failure in the face of a natural hardship. In our case, extinction would be due to a uniquely different reason: inordinate success."

9.5 Consciousness of Electrochemical Machines

The human brain is the most mysterious natural organ, which can be approached from several divergent points of view, from electrochemistry to psychoanalysis. This is where ancient ideas clash: Is the mind a ghost in the machine? Is there such a thing as the soul? What is consciousness? Is there such a thing as free will? On the other hand, if this is just a machine, how does it work? Research is digging its tunnels from two sides: elementary mechanisms or the holistic response to external impulses. The tunnels are multiple, but it is not clear where or when or whether they will ever converge.

From the elementary side it looks more or less clear. The brain is a collection of interconnected nerve cells, just as a computer's CPU is a collection of interconnected transistors. Nerve cells differ from other kinds of cells by their dendritic structure, sending out long axons ending at synapses where they contact other nerve cells, as shown in Fig. 9.15. Omitting fine detail, it works by plain electrochemistry and ion transport (Hodgkin and Huxley, 1952), as schematized in the lower right-hand panel of the same figure. In a resting neuron, the potential is negative inside

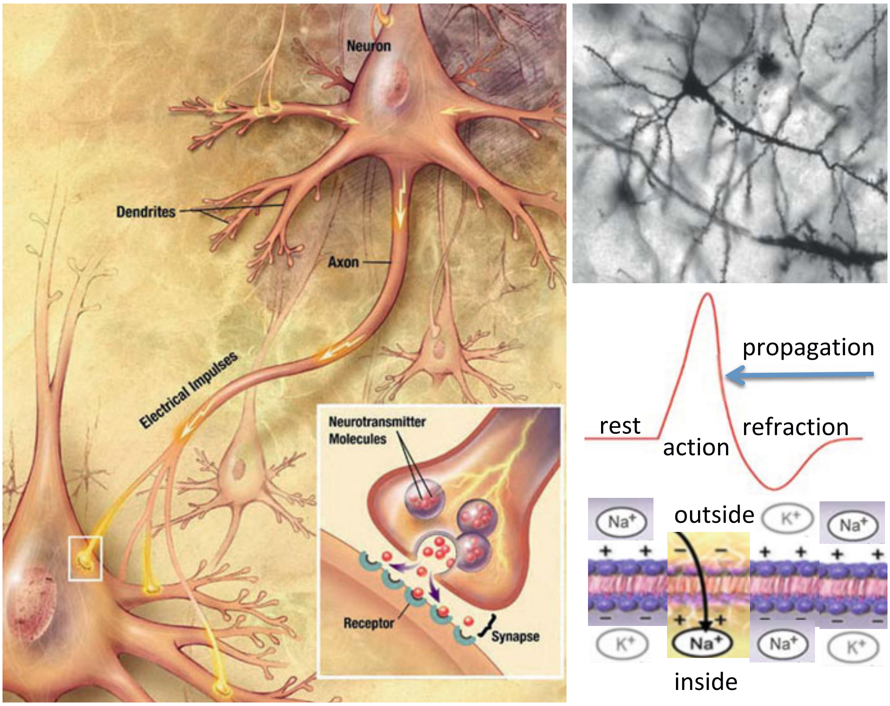


Fig. 9.15 *Left*: A neuron, with a synapse in the *inset*. *Top right*: Connected neurons in the brain. *Bottom right*: Propagation of an impulse through an axon (see the text for explanation)

the cell because of an excess of sodium ions Na^+ on the outside and potassium ions K^+ on the inside. When the action potential is triggered (the neuron is "fired") in response to synaptic signals, the polarization reverses, Na^+ channels in the membrane open, and Na^+ ions rush in. This stimulates neighboring Na^+ gates to open, and in this way the action potential travels down the length of the axon to a synapse where it signals other neurons. In the wake of the propagating pulse, the distribution of K^+ and Na^+ is reversed and the neuron enters a refractive phase, still more negative on the inside and unable to react to a new excitation; now K^+ channels open and the original state is restored.

The Hodgkin–Huxley equations describing this process are already able to initiate complex dynamics; we saw in Sect. 7.2 that even their simplified version, the FitzHugh–Nagumo equation, generates a variety of patterns. Take two neurons modeled in this way and you can already have chaos. But the complexity of the human brain is striking by the sheer numbers involved: there are a hundred billion neurons, each connected on average with ten thousand others. Electronic complexity is multiplied by chemical complexity: axon potentials and also impulse transfer at synapses all depend on zillions of chemical agents circulating in the body. The darling of the neuroscientists is the worm *Caenorhabditis elegans*, which sports 302 neurons, with their connections fully mapped – but this is still a complex network and its behavior is unpredictable.

In the human brain, establishing neural connections is an epigenetic process. Although intellectual abilities are inheritable to some extent, the connectivity of our brains develops throughout our lives, most intensely during childhood, influenced by education, and declining in old age. In this way, every person is self-built and unique. The Nobel prizewinner Gerald Edelman (1987) put forward the idea of "neural Darwinism" – natural selection among developing neural circuits. This, if true, follows the rules of all complex networks. Google's algorithm works in a similar way – but brain connections are material and rewiring cannot be easy. Memories are accumulated in these circuits, perhaps in distributed "holographic" ways.

Looking at the brain from the outside, what can be said from a mechanistic point of view? Quoting Rabinovich et al (2006): "Specialized neurons transform environmental stimuli into a neural code. This encoded information travels along specific pathways to the brain or central nervous system composed of billions of nerve cells, where it is combined with other information. A decision to act on the incoming information then requires the generation of a different motor instruction set to produce the properly timed muscle activity we recognize as behavior." Is this any different from the operation room of a chemical factory or a self-driven car? I hope we are more than this. "I am not a robot", as we often declare to a suspicious web site. The epigraph to the cited long review is taken from Jacques Hadamard, a mathematician thinking about psychology: "Will it ever happen that mathematicians will know enough about the physiology of the brain, and neurophysiologists enough of mathematical discovery, for efficient cooperation to be possible?" This can be paraphrased more modestly: "Will it ever happen that mathematicians will know enough about brain interconnections and neurophysiologists enough about the physiology of the brain?"

The combined billion euro European Human Brain Project (HBP) and the US BRAIN initiative aspire to unveil brain wiring through imaging and simulation. The "neuromorphic" computing system in Heidelberg mimics two hundred thousand neurons and fifty million synapses on a silicon wafer. Neurophysiologists have mapped different functions of the various regions of the brain – which are still macroscopic, containing a huge number of neurons. HBP trumps this with a Big Data attack combining tens of thousands of different imaging studies with extensive cognitive tests. Marcello Massimini claims on the HBP web site that "we are close to a profound leap in our understanding of consciousness, like before Charles Darwin began mining the rich seams of evidence that led to the theory of evolution". For sure, there is no comparison between the HBP data bank and the meager evidence Darwin had at his disposal – but it is still all about wiring, anatomy, and electronics.

Experimental neuroscientsts view the brain in a different way, but still from the outside. "Brain secretes mind as the liver secrets bile", said Pierre Jean George Cabanis more than two hundred years ago. Why would it not be subject to chemical inputs as the liver is? Even the Hodgkin–Huxley model admits this. In the bestseller hailed by reviewers as one of the best nonfiction books they have ever read, Robert Sapolsky (2017) describes the straightforward chemistry of manipulating emotions by hormones and drugs. This should be more efficient than talking to a shrink, but personally, it gives me an uneasy feeling. We are squeezed between great efforts to describe ourselves as sophisticated electronic circuits, on the one side, and chemical machines, on the other. Both sides are right in their limited vision, and both are helping mentally and sensorily impaired people in different ways. Perhaps electronics and chemistry will even come together at some point – but we are lucky that these breakthroughs did not come earlier, when the 20th century dictators were around. Criminal lawyers are now turning to neuroscience, and one judge has already allowed a scan showing a murderer's brain activity to be admitted as evidence to support a claim of innocence (Davis, 2017). If a machine kills you, it is not to blame, it's somebody else's fault.

But what about understanding consciousness? Is it just a hype to satisfy funding agencies? What about free will? If my decisions are governed by an interplay of electric pulses modified by hormones and affected by the intrinsic chaos of an unfathomably complex network, am I just a "reactive robot, slave to natural law" as some would say? This interplay *is* the conscious self, and my behavior stems from processes within *my* unique brain, with all the electrochemistry and connectivity which are exclusively *mine*. As long as I am not drugged or tortured, what this machine decides is *my* decision. No dualism is involved here: the machine is the ghost. Quantum uncertainty may or may not play a role in mental processes: classical uncertainty of a complex system is enough for unpredictability. The soul is an emergent entity: the matter that had been traditionally viewed as inert ceases to be such when it is complexly organized. Francis Crick (1994) said: "You are nothing but a pack of neurons." We are, but we are more than this because we are uniquely made and interconnected, very big, and very finely structured packs.

The brain is no more a Turing machine than an embryo is a Turing pattern. The complicated chemistry affecting transmission of signals and synaptic contacts ensures that its operation is both digital and analog, if we are allowed to use computational terms, and a qualitative transition in complexity occurs somewhere on the way from the *C. Elegance* worm to the human. The analog chemistry, exploited by neurochemical drugs, is not uploaded to the Heidelberg machine. Not that I would object in any way to the HBP project. A billion euros could be spent far more wastefully than giving food for thought to smart people, and even literally feeding them, while they observe intricate phenomena and hide from bureaucrats under a screen of hype.

9.6 Environment

This chapter started with turbulence and ends with turbulence on a larger spatiotemporal scale, and turbulence not restricted to mere hydrodynamics but extended to interactions between sunlight, atmospheric composition, oceans, vegetation, and the bowels of the Earth. Everybody is now talking about the problems of defending an

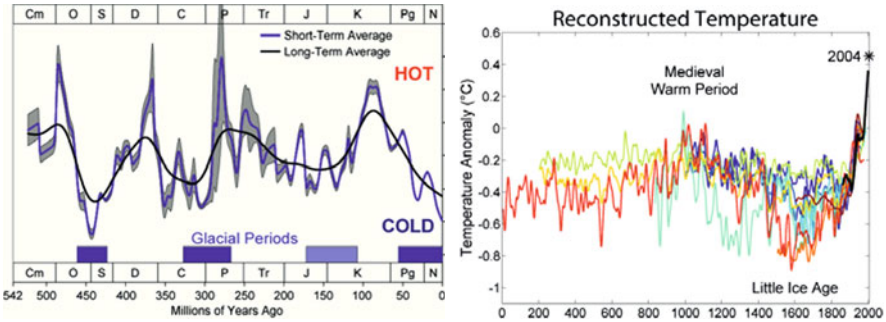


Fig. 9.16 *Left*: Change in global temperature during the past half a billion years. *Right*: Temperature changes in Europe over the last two thousand years according to different estimates

environment that is suffering from the activities of a single species of ape, endangering itself by its own great success.

This species evolved on a planet that was not distinguished by stability and permanence. During the hundreds of millions of years of the evolution of life, it passed through "snowball Earth" periods of global glaciation interspersed with a prevalence of wet jungles (Fig. 9.16). Continents moved and massive eruptions and meteorite impacts contributed to mass extinctions. The species I have in mind has benefitted from this turmoil, as meek mammals inherited an Earth swept clear of dinosaurs; its most significant mental and cultural development took place during the last glacial period. But now that we are in a domineering position, we don't want drastic changes anymore.

Joseph Fourier (1827) was the first to propose what is now called the greenhouse effect, warming due to absorption by atmospheric "greenhouse gases" of infrared radiation emanating from the Earth. The strongest effect comes from water vapor but it cannot be blamed, as it is ubiquitous and changeable, so the main culprit is carbon dioxide, CO_2, whose historic fluctuations in the atmosphere correlate with fluctuations in temperature. A more potent greenhouse gas is methane, which is apt to be involved in a positive feedback loop: as the temperature rises, methane is released from permafrost and this multiplies the effect. Changes in albedo, the ratio of the irradiance reflected to the irradiance received, are also potentially destabilizing: albedo goes down as snow melts. No wonder the climate fluctuated so wildly in the world where we happen to live, with such positive feedbacks running loose.

Glaciers covered what are now the most prosperous countries only about fifteen thousand years ago, and it is suspected that we are living in an interglacial period with an unstable climate. The medieval warm period when the Vikings traveled to Greenland and Newfoundland gave way to the "little ice age", familiar from the sight of skaters on frozen channels, as recorded by Dutch painters. Walther Nernst and Svante Arrhenius, laureates of the Nobel Prize in Chemistry who were wary of a new ice age, suggested setting fire to unused coal seams to increase the global temperature. Their advice has been followed spontaneously in some sense. What is being burned in increasing quantities is oil rather than coal, and rising temperatures have brought about a reverse concern. Carbon dioxide emissions are now called "poisonous", although this "poison" is what plants live on. Really poisonous sulphurous emissions blocking sunlight have been largely eliminated by advanced technologies, and this has also contributed to the temperature rise in the late 20th century. There are proposals to inject sulphur particles into the upper atmosphere, where they would do their job without poisoning us. Like all global engineering projects, this is dangerous even from a legal point of view: entire countries may sue for damages, so it remains to wait for a major volcanic eruption to counterbalance the greenhouse effect, just as it might have triggered the "little ice age".

On the web site of O'Reilly Media, Tim O'Reilly compares the climate change dilemma with Pascal's wager. In both cases, we play a game against God or Nature. In the spirit of game theory, not yet invented, Blaise Pascal compiled a risk matrix. If God exists and we believe in Him, we gain eternal bliss, and if we don't believe, our loss is infinite as well, while other matrix entries are finite. The conclusion is clear, but somehow non-believers are unlikely to be persuaded. O'Reilly follows suit: "On one side, the worst outcome is that we've built a more robust economy. On the other side, the worst outcome really is hell." He does not say whether he believes in heaven and hell. More precise versions of the dilemma have compared the economic outcomes of different strategies for fighting global warming, weighing future gains against current expenditures. The proposal by Al Gore was judged by far the most expensive. Thankfully, he was given a Nobel Peace Prize (2007) and not the Physics Prize for his powerpoint presentation on climate change. Doing nothing turned out to be one of the best strategies from the economic point of view. Studies of this kind, replacing infinities by terabucks, cannot be relied upon either: future losses are discounted against current expenditures at some interest rate, and this greatly affects the results; and how shall we price human suffering?

Climate models are no more reliable than economic ones. They disagree among themselves, though they all show a warming trend, which is probably to be expected, because their basic assumptions and methodology are similar. Much of the uncertainty comes from the most important component of the Earth environment – water. Water vapor is a potent greenhouse gas; on the other hand, clouds have a high albedo and the oceans absorb CO_2. The best way to test the models would be to run them from the past comparing with real data, but precise records exist only from the mid-20th century, which is too short an interval, and paleoclimatic data lack precision and detail.

Some suggested ways to reduce anthropomorphic influences on climate change, like developing green technologies and renewable energy sources, are beneficial in their own right – but the problem is that, notwithstanding all media hype and international agreements, nothing essential is really being done. "Climate scientists" get their grants, dealers trade in carbon dioxide emission rights, and delegates to numerous conferences enhance CO_2 emissions by flying to their destinations. The debates become politicized: your opinion on global warming and ways of countering or adjusting to it would identify you as a Democrat or Republican – but neither side would give up driving cars and flying to a faraway vacation paradise. The newly affluent middle class in developing countries joins this celebration, while less fortunate people burn jungles, the "Earth's lungs" and a potent CO_2 absorber.

Economic growth is still the main measure of success of any government. Everything tends to grow exponentially, both production and consumption. The old tale of a wise man arranging to be paid by placing a single grain on one square of a chess board, two on the next, and so on, till the entire wealth of the kingdom is exhausted, should have taught humanity the dangers of exponential growth. Two percent a year is not as evident as doubling but it comes to the same thing in the long run. The capitalist economy needs the profits brought by its expansion. So far, this has worked, and even oil reserves, bolstered by the exploration of continental shelfs, have been growing faster than consumption. All this is bound to come to an end, but who really thinks on any time scale longer than a human life? For politicians, the most influential decision-makers, the time scale is the interval between two consecutive elections, for company managers, it is yearly reports and bonuses, and for stock traders, daily, if not millisecond, pricing fluctuations.

Thomas Malthus (1798) predicted famine as the inevitable consequence of population growth. This forecast has not materialized over the ensuing two hundred years thanks to great advances in agricultural technology, but this is still a short interval even on an historical, let alone evolutionary time scale. Population growth is not uniform. It is greatest where the economy is weakest, which is also caused by destabilizing feedback loops. Non-uniform development and mass migrations are not modern phenomena, but cruel negative feedbacks with the weak and unfortunate starving or exterminated by invaders are no longer operational. Migrations that once spread humans over the globe are now countered by political borders, but human waves persist.

Robert May, whom we already met in Sect. 8.2, now Lord May of Oxford and a member of the Committee on Climate Change, published a tongue-in-cheek prediction with two colleagues that *dragons* will appear due to global warming. The paper (Hamilton et al, 2015), complete with plots showing the high frequency with which dragons are mentioned in literature during the warm period of the 11th and 12th centuries, and its decline in the "little ice age" of the 15th to 17th centuries, followed by a recent sharp rise, predicts that we are rapidly reaching an optimum for breeding dragons, and recommends research into fireproof protective clothing. The authors take into account interactions between economic and environmental factors: "hoards that serve as homes to resting dragons are an ideal way to bolster a failing economic policy. This strategy of 'quantitative thieving' is highly likely to provoke reprisals from slumbering dragons who awake to discover that their nests [the mineral treasures of the planet?] have been stripped bare."

Chapter 10
Quo Vadis?

10.1 Where We Have Got to

When following the wide swings between complexity and simplicity, we may encounter simplicity through crude simplification but also by understanding roots and causes; and complexity through mix-ups and failure of analysis but also through interactions and enrichment. The simplicity we need most, that which the human intuition is able to comprehend, drives scientific insight; it is then dressed in refining details when applied to practical problems and branches out into its complex realizations, so that no way can be found to make it simple again. Like turbulent eddies cascading to shorter scales (Sect. 9.1), oscillations in our knowledge cascade down to the comprehension of details: any problem is at first hidden in a fog, which with luck will clear up – but finer features coming to light as time goes on will require further concentration and a magnifying lens to peer into the new unknown.

Unfortunately this refinement deprives us of the light of sudden comprehension; great aims disappear into an unreachable distance, like Kafka's castle. Where the ancient alchemist strove to find the philosopher's stone – which would be in our terms a catalyst for all reactions, both chemical and nuclear – the modern chemical engineer is happy to improve the performance of a particular kind of catalyst for a particular process. Where the elixir of life was sought, we are happy to find a drug reducing mortality in a single variant of cancer.

Expectations were already beginning to wither away at the dawn of the scientific revolution. Robert Boyle's wish list, prepared for the opening of the Royal Academy, but first exhibited publicly[1] in 2010, went far beyond the reality of the 17th century. However, it is highly specific and has now been largely fulfilled: "As you go down Boyle's list, some of the things sound quite silly, but then you realise we've kind of done them," – in the words of the Royal society librarian, as reported in the Guardian newspaper of 3 June 2010.

The list splits into several categories. The longest one is medicine. *The Prolongation of Life*: did he mean just extending life expectancy at birth, which at the time

[1] https://blogs.royalsociety.org/history-of-science/2010/08/27/robert-boyle-list/

L. Pismen, *The Swings of Science*,
https://doi.org/10.1007/978-3-319-99777-3_10

was less than 40 years, although much longer if childhood deaths and war casualties were excluded? If so, this has been achieved, even in excess, overburdening medical and social services. Very old people may feel and look like Jonathan Swift's struldbrugs, and some countries have already legalized euthanasia. *The Recovery of Youth* would be great, but will molecular medicine ever be able to achieve it? – and the continuation disappoints: *or at least some of the Marks of it, as new Teeth, new Hair colour'd as in youth* – these are just cosmetic improvements, and we have learned to care better for our teeth and even to grow or transplant new hair.

More medicine: *The Cure of Wounds at a Distance. The Cure of Diseases at a distance or at least by Transplantation*. Yes, we have this, but it's still better if the doctor is present while performing a heart transplant, rather than operating a robot by remote control. *The Attaining Gigantick Dimensions*. It is boring to think Boyle just meant an increase in average height by a few percent, which has little to do with scientific and technological progress. It is in fact our constructions and machines, which are attaining gigantic dimensions, so this wish should probably be moved to another category. Population numbers also keep increasing, but this is a cause for concern as well as gratification. *Potent Druggs to alter or Exalt Imagination, Waking, Memory, and other functions, and appease pain, procure innocent sleep, harmless dreams, etc. Freedom from Necessity of much Sleeping exemplify'd by the Operations of Tea and what happens in Mad-Men. Pleasing Dreams and physicall Exercises exemplify'd by the Egyptian Electuary and by the Fungus mentioned by the French Author. Great Strength and Agility of Body exemplify'd by that of Frantick Epileptick and Hystericall persons.*. This already smacks of illicit drugs and doping, so yes, we have that, but we would be better without it. Could it be that psychotropic drugs prescribed for any purposes, from healing to stimulation, are just as dangerous?

Under the heading of locomotion we find: *The Art of Flying.* Most probably, he had in mind human-powered craft, like the one sketched some hundred years earlier by Leonardo da Vinci, and imagined not once before that. Would the Montgolfier brothers with their balloon, still more than hundred years in the future, have qualified for mastering this art? Certainly, buckled up in reclining seats inside a dragon-sized jet, we ourselves would not qualify, while the aviation pioneers of the turn of the 20th century might. However, Boyle forgot to wish for an internal combustion engine to help us feeble humans. *The Art of Continuing long under water, and exercising functions freely there. The Emulating of Fish without Engines by Custome and Education only.* Here it is explicitly specified that engines are not allowed, so scuba diving might perhaps fulfill this wish – although it still requires an aqualung. A nuclear-powered submarine can continue under water almost indefinitely long, and besides, it can exercise functions that may wreck a civilization, but this involves more than just custom and education. *A Ship to saile with All Winds, and A Ship not to be Sunk.* Yes, we have ships that are now independent of the winds, but they can be sunk quite easily.

As regards chemical and biological engineering: *The Acceleration of the Production of things out of Seed.* We have indeed achieved great successes in increasing agricultural production, mostly through the chemistry of fertilizers and pesticides, but also through breeding; genetic engineering is taking its first steps, meeting social opposition. *The Transmutation of Metalls.* Boyle could hardly be satisfied by nuclear reactions transmuting single atoms. He was still an alchemist. *The making of Glass Malleable.* A singularly modest wish. *The Transmutation of Species in Mineralls, Animals, and Vegetables.* The first is accomplished by a multitude of chemical reactions. For the rest, genetic engineering again. *The Liquid Alkaest and Other dissolving Menstruums. Varnishes perfumable by Rubbing.* Modest wishes of a practising "Chymist" – unless we seriously include alkahest, a universal solvent sought by alchemists. Such a solvent should also dissolve its container – one of the problems encountered later with controlled fusion.

Under the heading of mechanical engineering: *The making of Parabolicall and Hyperbolicall Glasses.* We have huge telescopes, even in orbit, and multifocal glasses closer to the eye. *The making Armor light and extremely hard.* Anything, from flack jackets to tanks – but he does not wish for offensive weapons so extreme that they could break through any armour. *The practicable and certain way of finding Longitudes. The use of Pendulums at Sea and in Journeys, and the Application of it to watches.* This was the most urgent need, and the earliest to be fulfilled as a result of the high rewards established by the Longitude Act in the early 18th century. While the last wishes are the most realistic, this one is the most stunning: *A perpetuall Light.* Only the holy light of spiritual enlightenment, coming from God, can be perpetual. At least, so it was thought by Boyle's fellow believers. Unless what is meant is just electric light, but that would have been harder to foresee: if so, we have it, but it is only perpetual until blackout.

Boyle fared better than most science fiction writers. We have almost all of it and more, gained in small steps but driven by glorious or dangerous ideas. Some items are missing from the list: communication, entertainment, perhaps thought to be superfluous or not thought about at all. The list is utterly practical. It's all technology, rather than ideas.

We also have our wish lists, and in our time the future also both disappoints and brings surprises. Since the 1960s, fuel-cell-driven cars and fusion in plasma confined by magnetic fields have been expected to come within 20 years (the same as Nasreddin's donkey in Sect. 2.3). The former may be closer now, but regarding fusion, Pierre-Gilles de Gennes observed: "We say that we will put the Sun into a box. The idea is pretty. The problem is, we don't know how to make the box." Science fiction writers sometimes manage to foresee the near future in a modest way, but more often than not miss by a huge margin. In Stanley Kubrick's *2001: A Space Odyssey*, released in 1968, a passenger shuttle takes off to the Moon, as a matter of routine and with no security, just as the shuttle between New York and Washington once did – but a passenger uses a cabinet-sized contraption to make a video call. The connection between technology and social mores remains as unpredictable as ever, but the powers-that-be do seem to understand that it is science that brings golden eggs, and so they keep funding it, even supporting projects that will bring no economic or military gains.

10.2 Science Contemplates Society

Would Boyle have approved of the social changes that might bolster or accompany the technological developments he listed? Did he realize that the scientific ideas and methods that would make these advances possible might encroach on the Christian faith he cherished and the supremacy of the aristocracy to which he belonged? Why would he even think of that? He was an aristocrat, a good Christian, and immersed in scientific research, although his early years were violent times. Human affairs remained the prerogative of philosophers till "sociophysics" and "econophysics" joined biophysics in spreading the influence of physics far and wide.

Although attempts to describe social and economic phenomena in mathematical terms go back to the 19th century, and even as far as Daniel Bernoulli, they only really took off with the advent of Big Data in the late 20th century. The methods of statistical physics were extended to social phenomena in a straightforward way (Schweitzer, 2003), explaining, for example, the separation of social groups by neighborhoods and opinions in a manner similar to the onset of magnetization in the Ising model (Sect. 5.2). It is not difficult to extract a lot of data on opinions from social networks and get correlations, both evident and spurious. Detailed simulations are agent-based, and "sociophysicists" are well aware of the complexity of self-organized non-equilibrium systems and the effects of nonlinearity, fractal structure, and chaos that we discussed for far simpler examples in Chaps. 7 and 8. Juval Portugali, a prominent investigator of the structure and dynamics of modern cities, writes (Portugali, 2011): "There is nothing wrong of course in sophisticated simulation models crunching huge quantities of data by means of fast computers. What's wrong is, firstly, that simulation models originally designed as media by which to study phenomena of complexity and self-organization become the message itself." It is not difficult to tune parameters to reach desired conclusions, even unconsciously. A big city is as complex as a human brain, and there is no Hodgkin–Huxley model on the elementary level, as humans do not always behave rationally.

Chris Anderson (2008), the former editor-in-chief of Wired Magazine, proclaimed that "with enough data, the numbers speak for themselves", so that theory is obsolete: "correlation supersedes causation, and science can advance even without coherent models, unified theories". Indeed, mindlessly gathering data is much easier than developing theories. As a Russian saying quips, "let the horse think, he has a big head". Will our silicon horse give us a ride to deeper truths? The numbers are easily accessible, and are handy for correlating anything with anything else, but correlation is inferior to causation, as it does not distinguish cause from effect (Calude and Longo, 2017). T. S. Eliot presaged this before electronic computers arrived (Eliot, 1934):

> Where is the wisdom we have lost in knowledge?
> Where is the knowledge we have lost in information?

Is the new way of crunching Big Data a better predictor than the old way of discerning general laws, even qualitatively and imprecisely? Prophets have rarely succeeded, and if Nostradamus is often cited, it is because he took care to be vague and avoid mentioning dates and locations. Predictions by Karl Marx that capitalism would bury itself in its contradictions and in an extreme polarization of property and poverty have clearly failed. If there is a resurgence of polarization in the globalized society, where the winner can literally take all with a successful sales network, it is for another reason: Matthew's law, which we met in Sect. 5.3, describing the growth of big droplets at the expense of smaller ones.

If there are some laws of politics, society, and economics that are really working, they are qualitative and non-predictive in detail. The most universal one is the power law distribution. The Gaussian distribution that used to be beloved by statisticians emerges, like the Maxwell–Boltzmann distribution (Sect. 4.2), as a result of random fluctuations in large weakly interacting ensembles, with the average sustained in a natural way, by temperature in statistical physics or by biological laws setting the average human height. The power law prevails when these restrictions are absent. It appears in different disguises. Vilfredo Pareto (1916) quantified Matthew's law by digging the data on wealth in different countries at different times and finding that 80% of the wealth is invariably held by 20% of the population. These particular numbers are likely to change, as well as the exponent of the power law, and one can relate to them in different ways, approvingly, as Pareto did, seeing it as an expression of human nature, or, in contrast, by trying to redress it, as the welfare society does. The law is, however, even more universal. The distribution of city populations follows this law, which can also be attributed to Matthew's mechanism of riches bringing more riches – but Zipf's power law regarding the frequency of words in any language (Zipf, 1935) can hardly be attributed to competition among words: on the contrary, the more complex a speech element, the less frequently it occurs.

Another source of power laws is self-organized criticality (Bak, 1996). We met this phenomenon in Sect. 4.2 when talking about avalanches in granular media that are caused by the last grain overstepping the critical incline. Earthquakes happen when stress accumulates in faults to the point of overcoming friction, and their strength also follows a power law, with the strongest events being the rarest. Social upheavals may also be caused by accumulated stress, though not easily quantified, as revolutions ripen when "the bottoms don't want and the tops cannot live in the old way", as Lenin wrote in 1913. This is indeed what happened in Russia four years later, and in France in 1789, and in the 16th century Reformation triggered by excesses of art-loving Renaissance popes. Sometimes the overturn is engineered for the same reason from above, as in Russia at the turn of the 18th century and again in 1991. What is cruel in power laws is the low frequency and unpredictability of the strongest events, the Black Swans of Nassim Taleb (2007) and Dragon Kings of Didier Sornette (2003). While dragons sleep in their lairs (recall Robert May in Sect. 9.6), economists measure correlations and extrapolate trends, sometimes even getting Nobel Memorial Prizes (mistaken for the real Nobel Prizes) for their insights – until a major earthquake strikes.

After power laws comes the law of unintended consequences. This is a rather trivial effect of a failure while tinkering with a complex system that is only partly understood, as is the case for all complex systems. Take a trivial example: try to stop a frictionless pendulum at its equilibrium position – what could be easier? But suppose you are only able to observe its position from time to time, but not its velocity. So you see it deviating to the right and push it to the left, without knowing that it is going this way itself, and your interference only causes it to overshoot. What is the result? The pendulum will not stop, but change its regular oscillations to chaotic motion. On a more complex level, this effect of "ignorance begetting chaos" recurs in many attempts to regulate the economy and society. The stronger the regulation, the more disastrous the failure, although the effect may not be immediate, and this is one of the reasons why totalitarian regimes break down and empires fall.

Economy and financial markets may be the simplest of social phenomena, which, on the face of it, can be quantified, as claimed by numerous models by academic economists and traders. They are still immensely complex, not only due to unexpected external inputs but, most of all, due to internal interactions among a multitude of self-serving but irrational agents. The financial system has been built in a way the fox would build a chicken coop, to the dealer's advantage. It is developing toward faster operations of increasing complexity, with more sophisticated financial instruments and instant computer trading programmed by renegade physicists and mathematicians ("quants") trying to outdo each other in predicting market movements. The question which nobody appears to be asking is whether it serves in any way the production of goods which markets are supposed to regulate. Engineers designing control systems know a simple rule: the response time of a controller should be commensurate with the characteristic time scale of changes occurring in the system under control. Real material changes proceed at a far lower speed than market fluctuations, which generate nothing but nervous jitter, but benefit sophisticated players.

Sociophysics aspires to rescue us of dangerous uncertainties. The European-funded Visioneer project aimed to use "reality mining" to predict disasters and thereby avoid socio-economic crises and pathological collective behavior (Helbing and Balietti, 2011). The researchers are aware of the dangers of misinterpreting correlations, and look for more sophisticated forecasting tools based on the theory of complexity and chaos, such as a critical slowing down and intensifying correlations that could serve as advance warning signs. Didier Sornette (2003) promulgated the study of fluctuation spectra for prediction of both earthquakes and financial crises. The Visioneer recommendation to set up Crisis Observatories is destined to serve this purpose. As for climate change, "retrocasting" would help to hone the technique, but I am unaware of such studies. Nobody in the West anticipated the collapse of the USSR, but perhaps some analysis of kitchen chat in Moscow would reveal the causes, were it available. There is, however, plenty of information that could be used to "retrocast" the 2009 economic downturn and the 2011 "Arab Spring" descending into turmoil. The growth of fluctuations in anticipation of a crisis is questionable. Aftershocks always follow earthquakes, but foreshocks are weak and rare, and all social crises are followed by prolonged periods of unrest, while they may be preceded by dead calm.

Even if scientists managed to predict an impending crisis, it would be politicians that would have to implement the required measures. Unfortunately, our democratically elected politicians are often among the least qualified to understand scientific arguments. Though they often have a background in law or "political science" (if such a thing exists), they are becoming increasingly close to the entertainment "industry", which supplies the best tools for seducing the ignorant or indifferent sectors of the electorate that swing elections. Winston Churchill called democracy the worst system, apart from all the others, even though the British thanked him for winning the war by electing the non-entity Clement Attlee instead. Would scientists fare better? Plato's Republic was planned as a totalitarian state, and people would not tolerate even an enlightened tyranny. The ultimate obstacle to forecasts is the back-reaction of the complex conscious system of humankind, changing its behavior in response to the forecast. This is a source of classical tragedies: Oedipus would not have killed his father and married his mother if he had grown up in the family rather than been sent away in response to the prophesy.

Beyond all this, there are some existential problems that nobody, and least of all the present author, would dare to address. Humankind as a whole has inadvertently affected both environment and society, starting from the development of agriculture and going through industrial and informational revolutions, and there is no way the problems on a distant horizon can be predicted and avoided. Sometimes the obstacles are ulterior. Should we encroach upon the sanctity of human life? The great successes of medicine have prolonged the human lifespan, but they have brought about a great burden on a diminishing work force to support a graying population. Médecins Sans Frontières has succeeded in extinguishing diseases, and rising populations encroach on primeval forests and wildlife habitats. As old sexual prejudices fade away, women are not constrained by Victorian moral biases and take an equal part in social, cultural, and political life – but natality has dropped below the reproduction level in the enlightened countries, and is minimal among educated women. In the past, when child mortality, while extraordinarily high by modern standards, was lower among more successful population layers, downward mobility was observed, enriching the genetic pool. Contrariwise, a dysgenic downward trend in IQ tests is already being detected now. At the same time, countries retaining high birth rates unsupported by economic growth produce populations of unemployed and sexually frustrated young males, feeding turmoil, extremism, and economic migration. The scientific approach stops here, giving way to ideals of humanity which could only fail if a catastrophe came.

10.3 Society Shapes Science

Whether exploring Nature or society itself, science lives inside society and is subject to its laws. The essay mentioned in Sect. 1.6 starts on a high note (Ross, 1962): "The appellation scientist is considered a title of honor, hotly contended for by economists, engineers, physicians, psychologists, and others." The prestige of scientists ran high from the 1940s to the 1960s, but the reason for this was not so honorable: nuclear bombs and rockets. In a survey of Soviet high school students in the 1960s, physicist was almost the top-ranking profession, second only to cosmonaut. Stalin doubled scientists' salaries near the end of World War II, bringing a common PhD (Russian "candidate") to the level of the Soviet upper middle class (which would still be below the poverty line in the US). The Soviet elites – full members of the Academy of Sciences, top artists, top sportsmen, approved authors – enjoyed about equal levels of renumeration and perks; just compare it with the wide gap between star scientists and other kinds of stars all over the world these days, not to mention the financial elites. The Soviet governing elite also trumped everybody else – but they observed that Western top elites do better; this was the true cause of the collapse of communism in 1989–1991.

The social prestige of science dwindled after the last man left the Moon and *détente* shifted attention from rockets to consumer goods. A scientist's work is full of dreary routine tasks. Capable PhDs with the good fortune to get tenure track positions have hardly any time to do their own research; besides teaching, the schedule is filled up by writing research proposals and reports, struggling with journal referees to squeeze in publications, and refereeing submissions by others going through the same ordeal. The data-driven approach extends to hirings and promotions, weighing up not only the number (rather than quality) of publications, but also the ranks of journals in which the candidate publishes. Scientific and near-scientific journals thrive on this supply. Materially, publishing has never been easier than now, when authors provide formatted "manuscripts" (which no hand has ever touched), referees are not paid, and editors do not edit, and the spreading hordes of "open access" journals have little to care about but maintaining web sites and collecting publication fees. Publications are also becoming more and more multi-authored. On the one hand, this is encouraged by the ease of communication, enabling authors at opposites sides of the world to write and rewrite a shared text, while on the other, it is driven by increasing funding demands and the need for sophisticated specialized tools and a wide range of narrow areas of expertise.

Funding agencies have built up huge bureaucracies. Thus, the European Research Council, a generous monster, requests its beneficiaries to report every cent and every minute spent on any budget item by filling out a cornucopia of forms. These bureaucracies are copied within universities, thereby cutting a lion's share of overheads from hard-won grants. The actual work is left to graduate students and postdocs fed by the same grants; if some of them are lucky enough to acquire scientific skills and promotion, they enter the same vicious circle, and most often continue the same work they were doing under supervision but at a further level of detail. Promoting the hand-cranked calculating machine he had invented in 1673, Gottfried Wilhelm Leibniz said: "It is unworthy of excellent men to lose hours like slaves in the labour of calculation which could safely be relegated to anyone else if machines were used." Alas, we waste more hours like slaves with our much better computers.

Scholastic disputation was one of the main tools of academic discourse when it prospered in ancient Greece and in Europe in the high middle ages. It was often ridiculed by civil society. Aristophanes famously cartooned Socrates; neither medieval scholars were held in high esteem. As Donna Bianca said in a Heinrich Heine's poem, "Capuchins and rabbis stink the same". Ridiculous as it sometimes was, the disputation habit has died away in our day. Scientists give seminar talks:

45–50 minutes with 5–10 minutes for polite questions. At large conferences, it may be 12–15 minutes including discussion, rising to half an hour at small specialized workshops. If you insist with a particularly apposite query, the chairman will cut you short, and your neighbors will look with disapproval at a spoilsport delaying the coffee break, the time set aside for leisurely and amiable chit-chat. If you come out with a talk containing long formulas which are hard for the audience to understand, you will lose any chance of being invited next time. It is much better to have lots of colorful pictures, or still better, movies, and who will be able to check how they were cobbled together? This is good for publications, too. Even serious journals adorn tables of contents with "graphical abstracts", apparently in case some illiterate user should chose to glance through them. Thankfully, not only representations of Big Data but microscopic structures, mesoscopic patterns, and realistic simulations look well on the screen.

Only outstanding claims, such as the observation of polymerized water by Boris Derjaguin (the leading figure in colloid chemistry and surface phenomena) in the 1960s (Derjaguin and Churaev, 1973), or cold fusion, a nuclear reaction in a test tube (Taubes, 1993) by Martin Fleischmann (a leading electrochemist) in 1989, are able to stir up any audible controversy. Lesser false findings are quietly buried in scientific journals, and even keep getting cited; only in biology and medicine, where they might have dire consequences, are they sometimes withdrawn with an accompanying scandal. A paper stating that most claimed research findings in this field are false and may often be simply accurate measures of the prevailing bias (Ioannidis, 2005) received about three thousand citations. A study by the Open Science Collaboration (Aarts et al, 2015) found only 36 percent of the findings published in top-rank psychology journals to be successfully replicated. A model of cohorts of diligent or careless or unethical scientists with funding allocated according to their published output (Grimes et al, 2018) suggested that the trustworthiness of published science is influenced by pressures for positive results encouraging false reports. This is an old story. Richard Feynman (1999), without the benefit of extended statistics, talked about it in his 1974 commencement speech. The preceding chapters are full of mix-ups and errors made by famous scientists, and even Newton and Einstein made mistakes – but it becomes far worse when funding pressure drives researchers to imitate the methods used by advertisers.

The creators of quantum mechanics clashed in their disputations in Copenhagen and Göttingen. This tradition was carried by Lev Landau to the Soviet Union, where it flowered among the physicists in his circle, in spite of the general suppression of discourse. The tradition of Landau's seminars was spread around by his numerous brilliant students. One of them was my mentor, Benjamin Levich, once chided by Landau for dropping out of "real" physics to work on "easy" physicochemical problems. This kind of seminar was like an intellectual rugby match. The sport was to attack the speaker from all sides, and the seminar continued till the audience was persuaded to their satisfaction (or exhausted), and sometimes reconvened the following week after four hours of indecisive mayhem. Nobody was offended, as a rugby player wouldn't grudge a jolt in the heat of a game. Some emigrés, failing to shed their habits in the civilized West, had trouble with getting tenure, notwithstanding their skills; I knew a talented but vitriolic scholar who committed suicide after being fired in both Israel and the USA.

Occasionally, discussion still makes its way into academic journals. The leading physical journal, Physical Review Letters (PRL), sometimes publishes comments criticizing a recent publication, together with its authors' response, but it does this rather reluctantly, setting strict but somewhat arbitrary rules about what constitutes a comment. Moreover, sending a comment is not a prudent decision: you can only make enemies, and who knows what surprise they may prepare for you. Being a run-of-the-mill professor, I generally play by the rules, but once (in fact in 1994), together with a friend and coauthor, we published a comment criticizing a paper we judged to be utterly wrong. Our paper on a related subject, submitted to PRL shortly afterwards, was rejected, I suspect, following a review by the same person that we had so unwisely excoriated. It was published, of course, in a lesser journal, as all rejected papers are, and I cherish it as one of my best, even though it is scarcely ever cited.

I repeated this mistake only once, quite recently (2016), being too angry at what I read, and argued with an editor (perhaps an author's friend) for months. The comment was eventually rejected on formal grounds, but upon my appeal, a more highly placed editor (perhaps my friend) ordered a normal reviewing procedure to be carried out, so that the comment was published almost a year after it was submitted. Such stories are probably not uncommon, but do not come to the light of day. It can be worse: two of my colleagues noticed in the same prestigious journal (dedicated to publishing innovative work of general interest to physicists) a paper exactly repeating their work of twenty years earlier. They sent a comment which was promptly rejected as not complying with the rules. One can understand, of course, the editors' reluctance to admit an editorial mistake or bias.

When I was sharing a house with Yves Pomeau in Santa Fe, we had the crazy idea of starting a journal dedicated to scientific discussions. Strangely, this idea was actually realized in some sense, as the Springer physics editor Christian Caron honed it down to a Discussions and Debates (DD) section of the European Physical Journal – Special Topics (EPJ ST). The general direction would be to discuss controversies arising in various branches of physics and controversial topics in technological development, also complementing it by criticism of the methodology used in certain research directions and reports of negative results, which may save the time and effort of other researchers pursuing the same line of investigation.

DD is not extinct at the time of writing, but is not particularly robust, as discussions are not more lively and penetrating than at scientific meetings. Our attempts to organize issues dedicated to controversial topics have failed. This was a preliminary subjective wish list: (1) Turbulence: Hierarchy of correlation functions *vs.* coherent structures. Can commercial fluid-dynamical programs be believed? (2) Superconductivity and superfluidity: Is the standard theory necessary to account for the experimental data? What is the nature of the supersolid state? (3) Glasses: Are they just metastable states of matter or another fancier equilibrium state? (4) Interfaces: The nature of line tension. The scaling at the contact line – how close can one approach? The role of surface roughness. (5) Biophysics: What are the practical limits of modeling? (6) Nanostructures: Laboratory curiosities *vs.* practical applications. (7) Quantum computing: Is it compatible with the foundations of quantum mechanics? Is it feasible? (8) Statistics: Effect of extreme events. Limits of non-extensive statistics. Limits of non-equilibrium thermodynamics.

The idea was to invite guest editors holding opposite views to follow a collision course in a single journal issue, but it did not work. In one attempt, a young professor believing that solar activity rather than anthropogenic influence is to blame for global warming tried to engage mainstream climate scientists in a discussion, but they would not budge to take part in a DD issue. An attempt to discuss the existence of supersolids, which is a comparatively "harmless" though nevertheless controversial topic, encountered the same problem. Two issues from the above list were published, rather circumscribed in their scope, plus a few others that were not too controversial, like the issue on "rogue waves", which worked out because there was little known on the subject and not yet any strongly competing schools of thought.

Stringent criticism finds its only refuge in anonymous reviews of journal submissions and, most importantly, grant applications. Science is a mass occupation nowadays, and scientists are common people who are not to blame for being wary of open discussion that could bring about undercover retribution, just as under totalitarian regimes citizens would not speak their mind lest they were reported by a snitch. Isn't this attitude common in society at large? In the Soviet Union of my youth, we would certainly not have opened up our souls in Facebook, had it then existed, or at a trade union meeting, but we would say whatever we wished in the kitchen after transferring the telephone (which was believed to be tapped even when idle) to the balcony.

Supervision by security services is replaced by the pressure to be "politically correct" in the free democratic West. Some questions, which are scientific on the face of it, like climate change, have become politicized. Race and gender issues are still more highly charged. Observe that an Olympic short-track final is filled up exclusively by descendants of West African tribes, while athletes from East African highlands dominate long distances and Europeans are doing well in the middle range. Is this remark not racist, by any chance? Discussing "nature *vs.* nurture" may be as slippery as walking on ice. Racial differences in GRE scores are attributed (Miller, 2013) to social factors that should be redressed by affirmative action, but nobody would suggest setting up a quota for white and Asian basketball players in top teams. Is it sexist to have separate sport competitions for women, even in chess, or do we need to strive for gender parity in sports, as well as in physics departments? Women are better than men in many ways, and top female athletes would outrun the overwhelming majority of men, even though they consistently show results about 10% lower than their top male counterparts. Equality of civil rights and prospects does not mean that all people are equal in all respects.

10.4 The End of Science?

In spite of all the pressures at large in society, old-style science sometimes makes a comeback. Alex Müller labored with no funding to discover high-temperature superconductivity in ceramics and win the 1987 Nobel Prize in Physics. Andre Geim and Konstantin Novoselov used no more costly equipment than scotch tape to isolate single-layer carbon sheets of graphene, which brought them the 2010 Nobel Prize. Science, though less well remunerated and leaving less free time than many other "white collar" (often replaced now by T-shirts) occupations, brings all the rewards of continuing intellectual progress. I was distressed when my talented postdoc, frustrated by the prospect of laboring through the German system, defected to a start-up training bots to increase the profits of insurers – at the expense of the insured, who else? How many more minds are wasted in "technology" companies toiling on an umpteenth version of an app, only to bother us with "upgrading" it again, and often only to slow it down? Any middling research work contributes more to society than this. How many brilliant minds are attracted to sterile but lucrative financial games?

As the first millennium was approaching, Christians all over medieval Europe were preparing for the End of the World. On the eve of the second millennium, the end of science was prophesied, though with less awe (Horgan, 1996; Lindley, 1993), alongside the end of history, announced prematurely by Francis Fukuyama (1992) and antedated by Günther Stent (1969). John Horgan's book caused quite a stir, disavowed by some prominent scientists, who counteracted by asserting the acceleration of scientific research into a broad range of unsolved problems – but it was certainly gratifying to those wishing to "de-center" science as "just one amongst a plurality of ways-of-knowing the world" (Cunningham and Williams, 1993). The divide between "accelerationists" and end-of-science adepts continues, and prompts discussion of the reasons why the contemporary scientific worldview has made the emergence of this problem possible (Shkliarevsky, 2013).

Debates of this kind are full of uncertainties. As one cannot see into the future[2], the discussion may only rely on general principles, which can be interpreted in different ways. The analogy with Earth exploration suggests that the period of great discoveries ends when the limits are reached. Just as Magellan's ship, circumventing the globe, could not continue to the Moon, we are barred by the laws we have discovered from superluminal travel, and other universes, if they exist, are unattainable by definition, which is virtually equivalent to their non-existence. However, the Universe has not only width but depth. Circumventing the globe still left the jungles unexplored, not to mention the Earth's interior. As an area of knowledge expands, the boundary of exploration where it touches upon the unknown also gets bigger. Moreover, this boundary is multidimensional and fractal, and pockets of the unknown remain in the wake of progress.

But is it possible to measure discoveries? Can we weigh up the discovery of a huge black hole in the middle of a galaxy against the identification of an extinct

[2] According to a Danish saying used by Niels Bohr, and well worn since: "It is difficult to make predictions, especially about the future."

spider species embedded in amber? Some discoveries are more equal than others, and physicists, especially the physicists writing popular science books, treasure the extreme depths of the skies above and matter below more than Earth-bound inventions and curiosities. The output of “normal science” is growing exponentially, but where are the great breakthroughs? As the fictitious professor Mozart complained (Mermin, 1990): “All particle physics has taught us about the central mystery is that quantum mechanics still works. Perfectly, as far as anybody can tell. What a letdown! [...] For 65 years, since 1925 [the birth of quantum mechanics], we’ve been probing, at finer and finer levels. That’s more than a quarter [now more than a third] of the time between 1685 [Newton’s *Principia*] and 1925. And more of us have been working on the problem than the world’s entire supply of physicists between Newton and Bohr. As for our funding, well, our funding has absolutely dwarfed all the combined funding from Bohr clear back to Archimedes. But what have we to show for it? Here we are today, another seven or eight orders of magnitude down beneath the level of the old revolution, and nothing fundamentally new is in sight.”

However, there has been great progress on length scales more relevant for human life, from nanometers to microns. Physicists have honed chemistry down to the point where they can build tiny structures on demand by operating on single atoms. Microscopy has been refined to atomic resolution, orders of magnitude below the light wavelength, through the use of single-photon interferometry. Nanorobots are preparing to detect and clean up budding cancer cells. The military are funding research on insect flight! Guess why? You may soon see (or maybe you have already seen but did not pay attention) a fly on the wall recording what you have just said and transmitting it to be decoded in a search for relevant information. And if somebody needs it badly enough, another fly may land on your bald spot and commit loud suicide. While some of these developments are still confined to the future, we can already date a movie with reasonable accuracy just by catching a glimpse of a phone or a computer on the screen. In the imaginary case of a nuclear holocaust and consequent partial recovery, archeologists will be puzzled by a mysterious sequence of shellac or vinyl shards, followed by scraps of tape with traces of magnetization, and then by silvery disks of unknown nature, all this followed by nothing at all, as if everything had flown up to the clouds, still below the radioactive layer. More likely, however, the layers will be intermixed, as the changes will have occurred too fast.

Science will not end while history continues:

> The endless cycle of idea and action,
> Endless invention, endless experiment,

seeking the knowledge – displacing the wisdom, as Eliot said, and before him Tennyson: “knowledge comes, but wisdom lingers”. Very likely, no more supercolliders will be built, but our eyes in the skies, looking out across all wavelengths with ever increasing resolution, will bring more news about the structure of the Universe, and even the deep structure of matter, both visible and dark, which wisdom unaided would not be able to comprehend. We may not travel to the stars, but we will better understand what is happening on this planet, from both artificial and living nanoscale structures to earthquakes – but probably not stock market crashes and

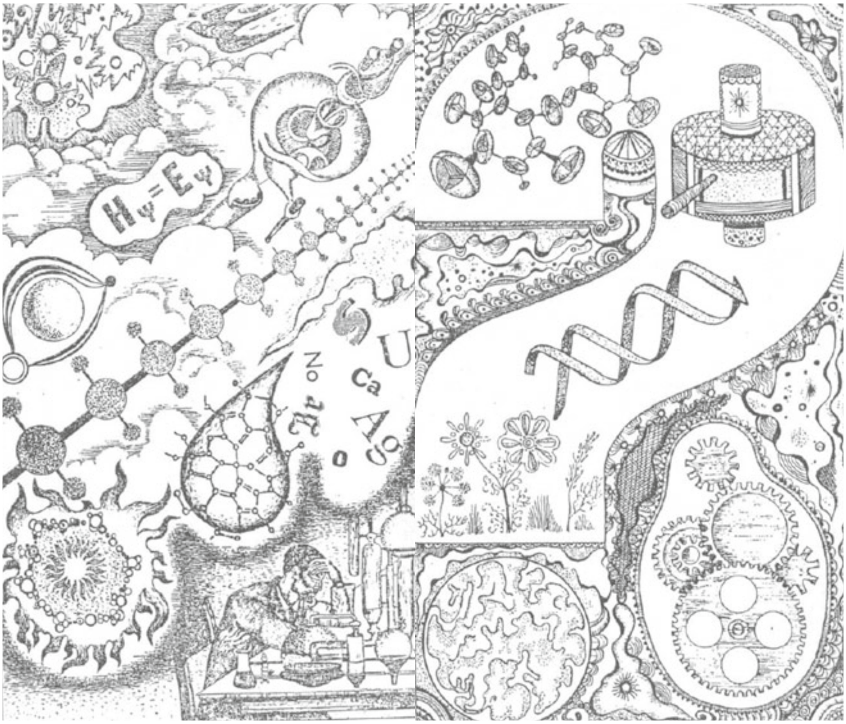


Fig. 10.1 The endless cycle of idea and action

bouts of depression, which, with sufficient wisdom, we might have avoided. Is there any chance that a new Newton or another Einstein will once again overturn our cherished paradigms? The newcomer may have trouble getting published, but will be able to deposit a paper on arxiv.org, provided this site remains online. Or could this great breakthrough be brought about by a hundred-strong CERN collaboration led by an Artificial Intellect? Or shall we just drown in the flood of information?

References

Aarts AA et al, 2015. Science, 349: aac4716
Abbott BP et al, 2017. Phys. Rev. Lett. 119: 161101
Abrikosov AA, 1957. Zh. Eksp. Teor. Fiz. 32: 1442–1452
Ahlers G, 1974. Phys. Rev. Lett. 33: 1185
Akhmediev N and Pelinovsky E, 2010. Eur. Phys. J. Special Topics 185: 1–4
Alexander JC and Antman SS, 1982. Quart. Appl. Math. 40: 83–92
Allen JF and Misener AD, 1938. Nature 3558 # 141 p. 75
Alon U, 2007. Nature Rev. Genet. 8: 450
Anderson C, 2008. https://www.wired.com/2008/06/pb-theory
Anderson R, 2016. https://www.edge.org/response-detail/26757
Andronov AA, Vitt AA, and Khaikin SE, 1937. *Theory of Oscillators* (In Russian, transl. Pergamon Press, 1966)
Aranson IS and Kramer L, 2002. Rev. Mod. Phys. 74: 99–143
Aref H, 2002. Phys. Fluids, 14: 1315–1325
Arnold VI, 1964. Dokl. Akad. Nauk SSSR 156: 9
Arrhenius SA, 1889. Z. Phys. Chem. 4: 96–116
Audoly B and Neukirch S, 2005. Phys. Rev. Lett. 95: 095505
Audoly B and Pomeau Y, 2010. *Elasticity and Geometry*, Oxford University Press
Avery OT, Macleod CM, and McCarty M, 1944. J. Exp. Med. 79: 137–158
Avogadro A, 1811. J. Phys. Chem. (Paris) 57: 58–65
Bacon F, 1620. *Novum Organum*, English trans. Clarendon, Oxford, 1878
Bak P, 1996. *How Nature Works: The Science of Self-Organized Criticality*, Copernicus, New York
Balmer J, 1885. Annalen der Physik, 3rd ser. 25, 80–87
Bardeen J, Cooper LN, and Schrieffer JR, 1957. Phys. Rev. 106: 162–164
Batchelor GK, 1967. *An Introduction to Fluid Mechanics*, Cambridge University Press
Becquerel H, 1896. C. R. Acad. Sci. Paris 122: 501–503
Becquerel H, 1900. C. R. Acad. Sci. Paris 130: 809–815
Bednorz JG and Müller KA, 1986. Z. Phys. B. 64: 189–193
Bell JS, 1964. Physics 1: 195.

L. Pismen, *The Swings of Science*,
https://doi.org/10.1007/978-3-319-99777-3

Bénard H, 1900. Ann. Chim. Phys. 7 (Ser. 23), 62
Ben-Zvi D and Barkai N, 2010. Proc. Nat. Acad. Sci. USA 107: 6924–6929
Berg CA, 2008. Tissue Eng. A 14: 1479
Bernal JD, 1954. *Science in History*, Watts & Co, London
Bernoulli D, 1738. *Hydrodynamica*, Argentorati, Strasbourg
Biberman LM et al, 1991. Usp. Fiz. Nauk 161 #6: 215–218
Biot MA, 1941. J. Appl. Phys 12: 155–164
Bloch F, 1928. Z. Phys. 52: 555–600
Bohm D, 1952. Phys. Rev. 85: 166
Bohr H, 1925. Acta Math. 45: 580
Bohr N, 1913. Phil. Mag. Ser. 6, 26: 476–502
Bohr N and Wheeler JA, 1939. Phys. Rev. 56: 426
Boltzmann L, 1897. Verhandlungen des 69. Versammlung Deutscher Naturforscher und Ärzte, Braunschweig, pp. 19–26
Bonn D et al, 2009. Rev. Mod. Phys. 81: 739–805
Borelli GA, 1680. *De motu animalium*, Rome
Born M, 1962. *Physics and Politics*, Oliver and Boyd, Edinburgh
Born M and von Kármán T, 1912. Phys. Zs. 13: 297–309
Boussinesq J, 1877. Mem. Acad. des Sci. Inst. Nat. France 23: 1–680
Boyle R, 1662. *A Defence of the Doctrine Touching the Spring and Weight of the Air*, Thomas Robinson, London
Brans CH and Dicke RH, 1961. Phys. Rev. 124: 925–935
Bravais A, 1850. J. École Polytech. 19: 1–128
Broglie L duc de, 1925. Ann. Phys. (Paris) Sér. 10, 3: 22
Buckmaster JD and Ludford GSS, 1982. *Lectures on Mathematical Combustion*, SIAM, Philadelphia
Burckhardt T, 1967. *Alchemy: Science of the Cosmos, Science of the Soul*. Trans. W Stoddart, Penguin, Baltimore
Burgers JM, 1940. Proc. Phys. Soc. 52: 23–33
Cahn JW and Hilliard JE, 1958. J. Chem. Phys. 28, 258
Calude CS and Longo G, 2017. Found. Sci. 22: 595–612
Carnot S, 1824. *Réflexions sur la Puissance Motrice du Feu*, Bachelier, Paris
Cartan É, 1922. C. R. Acad. Sci. Paris 174: 593–595
Cartwright ML and Littlewood JE, 1945. J. London Math. Soc. 20: 180–189
Cauchy AL, 1828. *Exercices de mathématique* 3: 213–236
Cavendish H, 1766. Phil. Trans. Roy. Soc. 56: 141–184
Chandrasekhar S, 1931. Phil. Mag. 11: 592–596
Chandrasekhar S, 1983. The *Mathematical Theory of Black Holes*, Clarendon, Oxford
Chen T et al, 2016. Curr. Opin. Chem. Biol. 34: 80–87
Chladni EFF, 1787. *Entdeckungen über die Theorie des Klanges*, Leipzig
Church AJ, 1879. *Stories from the Greek Tragedians*, Dodd, Mead & Co, New York
Clapeyron E, 1834. J. École Polytechnique 14: 153–190
Clausius R, 1850. Ann. Physik 79: 368–397, 500–524
Clausius R, 1865. Ann. Physik 125: 353–400

Clebsch RFA, 1864. *Theorie der Elasticität fester Körper*, BG Teubner, Leipzig
Copernicus N, 1543. *De revolutionibus orbium coelestium*, Johannes Petreius, Nuremberg
Couder Y and Fort E, 2006. Phys. Rev. Lett. 97: 154101
Coulomb C, 1784. Mém. Acad. Roy. Sci. 87: 229–269
Coulomb C, 1785. Mém. Acad. Roy. Sci. 88: 569–577
Crick F, 1994. *The Astonishing Hypothesis: The Scientific Search for the Soul*, Scribner, New York
Crick F, 1998. *What Mad Pursuit*, Basic Books, London
Cross MC and Hohenberg PC, 1993. Rev. Mod. Phys. 65: 851–1112
Cunningham A and Williams P, 1993. Brit. J. Hist. Sci. 26: 407–432
Curie P, Curie Mme P, and Bémont MG, 1898. C. R. Acad. Sci. Paris 127: 1215–1217
Dalton J, 1808. *A New System of Chemical Philosophy*, Bickenstaff, London
Darcy H, 1856. *Les fontaines publiques de la ville de Dijon*, Dalmont, Paris
Darrigol O, 2005. *The Worlds of Flow*, Oxford University Press
Davis K, 2017. *The Brain Defense*, Penguin Press, New York
De Duve C and Patterson N, 2010. *The Challenges of Genetics of Original Sin*, Yale University Press
De Gennes PG, 1971. J. Chem. Phys. 55: 572
de Gennes PG, 1975. C R. Hebd. Seances Acad. Sci. Ser. B 281: 101
de Gennes PG and Prost J, 1993. *The Physics of Liquid Crystals*, Oxford University Press
De Groot SR and Mazur P. (1962). Non-equilibrium Thermodynamics, North-Holland, Amsterdam
Derjaguin BV and Churaev NV, 1973. Nature 244: 430–431
Derjaguin BV, Churaev NV, and Muller VN, 1987. *Surface Forces*, Consultants Bureau, New York
Descartes R, 1637. *Discours de la Méthode*, Jan Maire, Leiden
Dirac PAM, 1928. Proc. Roy. Soc. A 117: 610
Dirac PAM, 1929. Proc. Roy. Soc. A 123: 714–733
Dirac PAM, 1931. Proc. Roy. Soc. A 133: 60–72
Doi M, 2013. *Soft Matter Physics*, Oxford University Press
Dyson FJ, 1972. Bull. Amer. Math. Soc. 78: 635–651
Dyson FJ, 1988. *Infinite in all Directions*, Harper & Row, New York
Eckert M, 2002. Arch. Hist. Exact Sci., 56: 451–468
Edelman G, 1987. *Neural Darwinism. The Theory of Neuronal Group Selection*, Basic Books, New York
Einstein A, 1905. Ann. Phys. (Berlin) 17: 891
Einstein A, 1905a. Ann. Phys. (Berlin) 322: 132–148
Einstein A, 1916. Ann. Phys. (Berlin) 49: 769–822
Einstein A, 1916a. Verh. Deut. Physik. Gesellschaft 18: 318–323
Einstein A and Infeld L, 1938. *The Evolution of Physics*, Cambridge University Press
Einstein A, Podolsky B, and Rosen N, 1935. Phys. Rev. 47: 777–780

Eliot TS, 1934. *The Rock*, Faber and Faber, London
Espinosa JR, Racco D, and Riotto A, 2018. Phys. Rev. Lett. 120: 121301
Ertl G, 1991. Science 254: 1750
Euler L, 1729. Comm. Acad. Sci. Petropoli 2: 347–368
Euler L, 1757. Mém. Acad. Sci. Berlin 11: 271–315 (trans. Physica D 237: 1825–1839, 2008)
Faraday M, 1831. Phil. Trans. Roy. Soc. London 121: 299–318
Faraday M, 1839. *Experimental Researches in Electricity*, Taylor and Francis, London
Farmer JD, Ott E, and Yorke JA, 1983. Physica D 7: 153–180
Feigenbaum MJ, 1978. J. Stat. Phys. 19: 25
Feynman RP, 1982. Int. J. Theor. Phys. 21: 467–488
Feynman RP, 1999. *The Pleasure of Finding Things Out*, Basic Books, New York
Fick A, 1855. Ann. Phys. 94: 59–86
FitzHugh R, 1961. Biophys. J. 1: 445–466
Floquet G, 1883. Ann. École Normale Supérieure 12: 47–88
Flory PJ, 1953. *Principles of Polymer Chemistry*, Cornell University Press
Fourier J, 1822. *Théorie analytique de la chaleur*, Firmin Didot Père et Fils, Paris
Fourier J, 1827. Mém. Acad. Roy. Sci. 7: 569–604
Frank-Kamenetsky DA, 1969. *Diffusion and Heat Transfer in Chemical Kinetics*, Plenum, New York (1st Russian edition 1947)
Fraunhofer J, 1814. Denkschr. Knig. Akad. Wiss. München 5: 193–226
Frenkel Y, 1939. Zh. Exp. Teor. Fiz. 9: 641
Frenkel Y, 1946. *Kinetic Theory of Liquids*, Clarendon Press, Oxford
Frenkel Y, 1946a. Usp. Fiz. Nauk 30: 11
Friedman A, 1922. Z. Physik. 10: 377–386
Frye N, 1957. *Anatomy of Criticism*, Princeton University Press
Fukuyama F, 1992. *The End of History and the Last Man*, Free Press, New York
Galileo G, 1623. *Il Saggiatore* (The Assayer), Rome
Gell-Mann M, 1962. Phys. Rev. 125: 1067
Gell-Mann M, 1964. Phys. Lett. 8: 214–215
Gell-Mann M, 1994. *The Quark and the Jaguar*, H Freeman, New York
Gibbs JW, 1876. Trans. Connecticut Acad. Arts Sci. 3: 108–248
Gibbs JW, 1902. *Elementary Principles in Statistical Mechanics*, Scribner, New York
Ginzburg VL and Landau LD, 1950. Zh. Eksp. Teor. Fiz. 20: 1064
Gödel K, 1931, Monatshefte Math. Phys. 38: 173–198
Goldman N et al, 2013. Nature 494: 77–80
Goldstein Sheldon, 1998. Phys. Today 51 #3 p. 42
Goldstein Sydney, 1969. Annu. Rev. Fluid Mech 1: 1–28
Golovin AA, Nir A, and Pismen LM, 1995. Ind. Eng. Chem. Res. 34: 3278–3288
Golovin AA, Nepomnyashchy AA, and Pismen LM, 2002. Int. J. Bifurc. Chaos 12: 2487–2500
Golovin AA and Pismen LM, 2004. Chaos 14: 845–854
Goodstein JR, 1984. Social Research 51: 691–708

Grimes DR, Bauch CT, and Ioannidis JPA, 2018. Roy. Soc. Open Sci. 5: 171511
Guth AH, 1981. Phys. Rev. D 23: 347–356
Hadamard J, 1898. J. Math. Pures Appl. 4: 27–73
Haeckel E, 1904. *Kunstformen der Natur*, Bibliographisches Institut, Leipzig
Hagberg A et al, 2006. Physica D 217: 186–192
Hamilton AJ, May RM, and Waters EK, 2015. Nature 520: 42–43
Harari YN, 2017. *Homo Deus: A Brief History of Tomorrow*, HarperCollins, New York
Harvey W, 1628. *Exercitatio Anatomica de Motu Cordis et Sanguinis*, Frankfurt
Hawking SW, 1974. Nature 248: 30–31
Hehl FW, Heyde P von der, and Kerlick GD, 1976. Rev. Mod Phys. 48: 393–416
Heisenberg W, 1925. Z. Physik 33: 879–893
Heisenberg W, 1948. Z. Physik 124: 628–657
Heitler W and London F, 1927. Z. Phys. 44: 455
Helbing D and Balietti S, 2011. Eur. Phys. J. ST, 195: 3–68
Helmholtz H von, 1858. J. reine und angewandte Math. 55: 25–55
Hertz H, 1893. *Electric Waves*, Dover Publications
Hessel JFC, 1831. *Krystallometrie, oder Krystallonomie und Krystallographie*, Leipzig
Hodgkin AL and Huxley AF, 1952. J. Physiol. London 117: 500–544
Hooke R, 1678. *De Potentia Restitutiva*, London
Horgan J, 1996. *The End of Science*, Basic Books, New York
Hubble E, 1929. Proc. Nat. Acad. Sci. 15: 168–173
Huh C and Scriven LE, 1971. J. Colloid Interface Sci. 35: 85
Hutton C, Shaw G, and Pearson R, 1809. Phil. Trans. Roy. Soc. London 43: 499
Huygens C, 1690. *Traité de la lumiere*, Pieter van der Aa, Leiden
Ioannidis JPA, 2005. PLOS Med. 2: e124 696–701
Ising E, 1925. Zs. Phys. 31: 253–258
Jaspers K, 1953. *The Origin and Goal of History*, transl. by M Bullock, Routledge & Kegan, London
Josephson BD, 1962. Phys. Lett. 1: 251–253
Joukowsky N, 1910. Zs. Flugtechnik und Motorluftschiffahrt 1: 281–284
Joule JP, 1845. Brit. Assoc. Rep., Chem. Sect. p. 31
Jülicher F et al, 2007, Phys. Rep. 449: 3–28
Kaiser D, 2011. *How the Hippies Saved Physics*, WW Norton & Co, New York
Kamerlingh Onnes H, 1911. Comm. Phys. Lab. Univ. Leiden # 124c
Kapitza P, 1938. Nature 3558, # 141, p. 74
Kármán, T von, 1921. Z . Angew. Math. Mech. 1: 233–252
Kekulé A, 1857. Ann. Chem. Pharm. 104: 129–150
Kekulé A, 1865. Bull. Soc. Chim. Paris. 3: 98–110
Kepler J, 1596. *Mysterium Cosmographicum*, Tübingen
Kepler J, 1609. *Astronomia nova*, Prag
Kerszberg M and Wolpert L, 2007. Cell 30: 205–209
Kibble TWB, 1976. J. Phys. A 9: 1387
Kirchhoff G, 1859. Monatsbericht König. Preuss. Akad. Wiss. Berlin, pp. 662–665

Kleman M, 1983. *Points, Lines and Walls*, Wiley, Chichester

Kolmogorov AN, 1941. Dokl. Acad. Nauk SSSR 30: 301–305

Kolmogorov AN, 1954. Dokl. Acad. Nauk SSSR 98: 527–530

Kröner E, 1966. *Theory of Crystal Defects*, Czech. Acad. Sci., Prague

Kuhn TS, 1962. *The Structure of Scientific Revolutions*, University of Chicago Press

Kuramoto Y and Tsuzuki T, 1976. Prog. Theor. Phys. 55: 356

Lagrange J-L, 1788. *Mécanique analytique*, la Veuve Desaint, Paris

Lamb H, 1895. *Hydrodynamics*, Cambridge University Press

Landau LD, 1937. Phys. Z. Sowjet. 11: 26–47, 545–555

Laplace PS, 1776. *Sur les ondes*, pp. 542–552

Laplace PS, 1798. *Méchanique Céleste*, Tome II, Livre IV, p. 543

Laplace PS, 1802. https://archive.org/details/philosophicaless00lapliala

Lavoisier AL, 1783. *Réflexions sur le phlogistique*, Académie des Sciences, Paris

Lavoisier AL, 1789. *Traité élémentaire de chimie*, Chez Cuchet, Paris

Leibniz GW, 1714. *Monadologie*, https://www.gutenberg.org/ebooks/39441

Leshansky AM and Pismen LM, 2010. Phys. Rev. E 82: 025301

Levich VG and Landau LD, 1942. Acta Physiochim. USSR, 17, 42–54

Levich VG, 1962. *Physicochemical Hydrodynamics*, Prentice-Hall, Englewood Cliffs, NJ

Lewis GN, 1916. J. Amer. Chem. Soc. 38: 762–786

Linde A, 2012. https://www.edge.org/response-detail/11981

Lindley D, 1993. *The End of Physics: The Myth of a Unified Theory*, Basic Books, New York

Lomonosov MV, 1760. *Meditationes de solido et fluido*, Academiae Petropoli

London F and London H, 1935. Proc. Roy. Soc. London A 149: 71

Lorenz EN, 1963. J. Atmos. Sci. 20: 130

Love AEH, 1892. *Mathematical Theory of Elasticity*, Cambridge University Press

Maddox J, 1999. *What Remains to Be Discovered*, Touchstone, New York

Madelung E, 1926. Naturwissenschaften. 14: 1004

Malthus TR, 1798. *An Essay on the Principle of Population*, J. Johnson, London

Mandelbrot BB, 1982. *The Fractal Geometry of Nature*, WH Freeman, San Francisco

Marangoni C, 1869. *On the expansion of a droplet of a liquid floating on the surface of another liquid*, Fratelli Fusi, Pavia

Marchetti MC et al, 2013, Rev. Mod. Phys. 85: 1143–1189

Maxwell JC, 1867. Phil. Trans. Roy. Soc, London 157: 49–88

Maxwell JC, 1873. *A Treatise on Electricity and Magnetism*, Clarendon, Oxford

May RM, 1976. Nature 261: 459–467

Mayr E, 1954. In Huxley J et al, *Evolution as a Process*, pp. 157-180, Allen and Unwin, London

McCurdy E, 1906. *Leonardo da Vinci's Notebooks*, p. 64, Duckworth, London

Meissner W and Ochsenfeld R, 1933. Naturwissenschaften 21: 787–788

Mendeleev D, 1871. J. Russ. Chem. Soc. (in Russian) 3: 25–56

Mermin ND, 1990. Phys. Today 43 #11: 9

Meron E, 2018. Annu. Rev. Cond. Mat. Phys. 9: 79–103

Michelson AA and Morley EW, 1887. Amer. J. Sci. 34: 333–345
Miller CW, 2013. APS News 22: #2
Misner CW, Thorne KS, and Zurek WH, 2009. Phys. Today 62: # 4 40-46
Moeendarbary E et al, 2013. Nature Materials, 12: 253–261
Mogilner A and Oster G, 2003, Biophys. J. 84: 1591–1605
Monine MI, Pismen LM, and Imbihl R, 2004. J. Chem. Phys. 121: 11332–11344
Mora S et al, 2010. Phys. Rev. Lett. 105: 214301
Moseley HGJ, 1913. Phil. Mag. Ser. 6, 26: 1024–1034
Navier CL, 1823. Mem. Acad. R. Sci. Paris 6: 389–416
Navier CL, 1827. Mem. Acad. R. Sci. Paris 7: 375–394
Ne'eman Y, 1961. Nucl. Phys. 26: 222
Nernst W, 1904. Z. Phys. Chem. 47: 52
Nernst W, 1918. Z. Electrochem. 24: 335
Neumann J von, 1951. In *Cerebral Mechanisms in Behavior* (LA Jeffress, ed.), pp. 1–31, John Wiley & Sons, New York
Newlands JAR, 1864. Chem. News. 10: 94–95
Newton I, 1686. *Philosophiae Naturalis Principia Mathematica*, S Pepys, London
Newton I, 1704. *Opticks*, S Smith, London
Nir A and Pismen LM, 1977. Chem. Eng. Sci. 32: 35–41
Noether E, 1918. Nachr. Ges. Wiss. Göttingen: Math-phys. Klasse, pp. 235–257
Normand C, Pomeau Y, and Velarde MG, 1977. Rev. Mod. Phys. 49: 581
Nusslein-Volhard C, 2008. *Coming to Life: How Genes Drive Development*, Kales Press
Onnes HK, 1911. Comm. Phys. Lab. Univ. Leiden, No. 122b, 1911
Onsager L, 1931. Phys. Rev. 37: 405–426
Onsager L, 1944. Phys. Rev. 65: 117–149
Ornstein LS and Zernike F, 1914. Proc. Acad. Sci. Amsterdam 17: 793–806
Oron A, Bankoff SG, and Davis SH, 1997. Rev. Mod. Phys. 69: 931
Osheroff DD et al, 1972. Phys. Rev. Lett. 28: 885–888
Ostwald W, 1900. Z. Phys. Chem. 35: 33
Ouyang Q and Swinney HL, 1991. Nature 352: 610
Panek R, 2011. *The 4 Percent Universe*, Houghton Muffin Harcourt, New York
Pareto V, 1916. *Trattato di sociologia generale*, G Barbera, Firenze
Patterson G, 2012. *A Prehistory of Polymer Science*, Springer, Heidelberg
Pasteur L, 1864. Conférences faite aux soirées scientifiques de la Sorbonne
Pauling L, 1931. J. Am. Chem. Soc. 53: 1367–1400
Pauling L and Wheland GW, 1933. J. Chem. Phys. 1: 362–374
Pelevin V, 2002, *Homo zapiens*, Penguin Books
Penrose R, 1960. Ann. Phys. (New York), 10: 171–201
Penrose R, 1969. Riv. Nuovo Cim. 1: 252–276
Penrose R, 2013. In *Computable Universe*, H Zenil (Ed.), World Scientific, Singapore
Perlmutter S et al, 1999. Astron. J. 517: 565–586
Pismen LM, 1973. *From the Miracle to the Number* (in Russian). Pedagogika, Moscow

Pismen LM, 1973a. Znanie Sila #2 p. 34 (in Russian)
Pismen LM, 1999. *Vortices in Nonlinear Fields*, Clarendon, Oxford
Pismen LM, 2006. *Patterns and Interfaces in Dissipative Dynamics*, Springer, Berlin
Pismen LM and Simakov DSA, 2011. Phys. Rev. E 84: 061917
Planck M (1900). Über eine Verbesserung der Wienschen Spektralgleichung. Verhandlungen der Deutschen Physikalischen Gesellschaft. 2: 202–204
Poincaré JH, 1890. Acta Math. 13: 1
Poisson SD, 1816. Mém. Acad. R. Sci. Inst. Fr., 2nd ser. 1: 70–186
Pomeau Y and Manneville P, 1980. Comm. Math. Phys. 74: 189–197
Popper K, 1959. *The Logic of Scientific Discovery*, Routledge, London
Portugali J, 2011. *Complexity, Cognition and the City*, Springer, Heidelberg
Prandtl L, 1904. Verh. Int. Math. Kongr., 3rd. Heidelberg, 1904, pp. 484–91
Priestley J, 1767. *The History and Present State of Electricity, with Original Experiments*, London
Prigogine I, 1955. *Introduction to Thermodynamics of Irreversible Processes*, Wiley Interscience, New York
Prout W, 1815. Ann. Phil. 6: 321–330
Rabinovich MI et al, 2006. Rev. Mod. Phys. 78: 1213–1265
Rayleigh, Lord and Ramsay W, 1894. Proc. Roy. Soc. London 57: 265–287
Rayleigh, Lord, 1916. Proc. Roy. Soc. London A 93: 148
Reinitzer F, 1888. Monatshefte für Chemie (Wien) 9 (1): 421–441
Richardson LF, 1961. Soc. Gen. Syst. Res. 6: 139–187
Richardson LF, 1922. *Weather Prediction by Numerical Process*, p. 66, Cambridge University Press
Riess A et al, 1998. Astron. J. 116: 1009–1038
Ross S, 1962. Ann. Sci. 18: 65–85
Rössler OE, 1976. Phys. Lett. A 57: 397
Rubio RG et al (Eds.), 2013. *Without Bounds: A Scientific Canvas of Nonlinearity and Complex Dynamics*, Springer, Heidelberg
Rumford, 1876. *The Complete Works of Count Rumford*, v. 5 pp. 469–493, Macmillan, Baths
Russell JS, 1834. Brit. Assoc. Adv. Sci. Report, pp. 531–544
Rutherford E, 1911. Phil. Mag. Ser. 6, 21: 669–688
Safran SA et al, 2005. Physica A 352: 171–201
Salbreux G et al, 2007. Phys. Biol. 4: 268–284
Sapolsky RM, 2017. *Behave*, Penguin Press, New York
Schamel D et al, 2014. ACS Nano, 8: 8794–8801
Schawlow A and Townes C, 1958. Phys. Rev. 112: 1940–1949
Schrödinger E, 1926. Ann. Phys. (Berlin) 384: 273–376
Schrödinger E, 1945. *What Is Life? The Physical Aspect of the Living Cell*, Cambridge University Press, Cambridge
Schwarzschild K, 1916. Sitzungsberichte König. Preuss. Akad. Wiss. 7: 189–196
Schweitzer F, 2003. *Brownian Agents and Active Particles: Collective Dynamics in the Natural and Social Sciences*, Springer, Berlin

Semenov N, 1935. *Chemical Kinetics and Chain Reaction*, Clarendon, Oxford
Serreli V et al, 2007. Nature 445: 523–527
Shannon CE, 1948. Bell System Tech. J. 27: 379–423
Shechtman D et al, 1984. Phys. Rev. Lett. 53: 1951–1953
Shkliarevsky G, 2013. Syst. Res. 30: 43–55
Sirovich L, 1987. Quart. Appl. Math. 45: 561
Sivashinsky GS, 1977. Acta Astron. 4: 1177
Slater JC, 1931. Phys. Rev. 37: 481–489
Smale S, 1967. Bull. Amer. Math. Soc. 73: 747–817
Sokal AD, 1996. Social Text, Summer/Autumn p. 216
Sornette D, 2003. Why stock markets crash, Princeton University Press
Stauffer D and Aharony A, 1992. *Introduction to Percolation Theory*, Taylor & Francis, London
Stent GS, 1969. *The Coming of the Golden Age*, Natural History Press, Garden City, NY
Stokes GG, 1845. Trans. Cambridge Phil. Soc. 8: 287–305
Stokes GG, 1880. *Mathematical and Physical Papers*, Volume I, pp. 197–229. Cambridge University Press
Stoney GJ, 1881. Phil. Mag. 5: 381–390
Su B-L et al (Eds.), 2012. *Hierarchically Structured Porous Materials*, Wiley-VCH, Weinheim
Swift JB and Hohenberg PC, 1977. Phys. Rev. A 15: 319
Tadmor Z and Gogos CG, 2006. *Principles of Polymer Processing*, Wiley-Interscience, Hoboken, NJ
Taleb NN, 2007. *The Black Swan*, Random House, New York
Tani I, 1969. Annu. Rev. Fluid Mech. 9: 87–111
Taubes G, 1993. *Bad Science: The Short Life and Weird Times of Cold Fusion*, Random House, New York
Taylor GI, 1923. Trans. Roy. Soc. London A 223: 289–343
Taylor GI, 1935. Trans. Roy. Soc. London A 151: 412–478
Tegmark M, 2014. *Our Mathematical Universe*, Alfred Knopf, New York
Thomson J, 1855. Phil. Mag. 10: 330–333
Thomson JJ, 1897. Phil. Mag. 44: 293–316
Thomson W, 1867. *On Vortex Atoms*, Proc. Roy. Soc. Edinburgh 6: 94–105
Thomson W, 1871. Phil. Mag., Ser. 4, 42: 448–452
't Hooft G, 2016, *The Cellular Automaton Interpretation of Quantum Mechanics*, Springer, Heidelberg
Timerman E and Timerman A, 2002. *The Elements*, Klaudpress, San Francisco
Truesdell CA, 1980. *The Tragicomical History of Thermodynamics 1822–1854*, Springer, New York
Turing AM, 1952. Phil. Trans. Roy. Soc. London B 237: 37–72
Tylor E, 1871. *Primitive Culture*, facsimile edition, Gordon Press, 1974
Tynianov YN, 1924. In *The History of Literature and Cinema* (in Russian), Nauka, Moscow (1977)
Uffink J, 2007. In *Handbook of the Philosophy of Science*, pp. 923–1074, Elsevier

Valadares, LF et al, 2010. Small 6: 565–572

van der Waals JD, 1873. *Over de Continuiteit van den Gas- en Vloeistoftoestand*, Sijthoff, Leiden

van der Waals JD, 1894. Z. Phys. Chem. (Leipzig) 13: 657

Vassiliev D, 2005. Ann. Phys. (Leipzig) 14: 231–252

Vicsek T et al, 1995. Phys. Rev. Lett. 75: 1226

Volovik GE, 1992. *Exotic Properties of Superfluid* ^{3}He, World Scientific

Volterra V, 1907. Ann. Sci. École Normale Supérieure 24: 401–517

Watson JD and Crick FH, 1953. Nature 171: 737–738

Weyl H, (1919). Ann. Physik, 59: 101–133

Weinberg S, 1989. Rev. Mod. Phys. 61: 1–22

Weinberg S, 2015. *To Explain the World*, HarperCollins, London

Wheeler JA, 1998. *Geons, Black Holes, and Quantum Foam*, WW Norton, New York

Whewell W, 1834. Quart. Rev. 51: 58–61

Wilczek F, 2012. https://www.edge.org/response-detail/10324

Wilczek F, 2012a. Phys. Rev. Lett. 109: 160401

Wilczek F, 2016. https://www.edge.org/conversation/frank_wilczek-power-over-nature

Wilson KG, 1971. Phys. Rev. B 4: 3174–3183

Wilson RW and Penzias AA, 1967. Science 156: 1100–1101

Winfree AT, 1984. J. Chem. Educ. 61: 661–663

Witten TA, 2007. Rev. Mod. Phys. 79: 643–675

Woit P, 2006. *Not Even Wrong*, Basic Books, New York

Wolfram S, 2002. *A New Kind of Science*, Wolfram Media, Champaign IL

Wollaston WH, 1802. Phil. Trans. Roy. Soc. 92: 365–380

Wolpert L, 1969. J. Theor. Biol. 25: 1–47

Wright R, 2009. *The Evolution of God*, Hachette, New York

Young T, 1802. Phil. Trans. Roy. Soc. London, 92: 12–48

Young T, 1805. Phils. Trans. Roy. Soc. London 95, 65

Zabusky NJ and Kruskal MD, 1965. Phys. Rev. Lett. 15: 240–243

Zaenen I et al, 1974. J. Molec. Biol. 86: 109–116

Zeldovich YB, Barenblatt GI, Librovich VB, and Makhviladze GM, 1985. *Mathematical Theory of Combustion and Explosion*, Consultants Bureau, New York

Zhabotinsky AM, 1964. Dokl. Akad. Nauk SSSR 157: 392

Zipf GK, 1935. *The Psychobiology of Language*, Houghton Mifflin, Boston

Zurek WH, 1985. Nature 317: 505

Zuse K, 1969. *Rechnender Raum*, Springer, Wiesbaden

Illustration Credits

Cover: Romanesque fresco of St. Proculus in the church of St. Proculus, 7th century, Naturns, South Tyrol, Italy. Dietrich Krieger CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0), via Wikimedia Commons
Figure 1.1: Public Domain,
https://commons.wikimedia.org/w/index.php?curid=466472
Figure 1.2 Illustration by V. Chernov (Pismen, 1973)
Figure 1.4: Courtesy TM/R Design
Figure 1.5: Illustration by V. Chernov (Pismen, 1973)
Figure 1.9(l): By Tetra quark – Own work, CC BY-SA 4.0,
https://commons.wikimedia.org/w/index.php?curid=38318435
Figure 1.11: By Joel Holdsworth – Non-Derived SVG of Radiate_gluon.png, originally the work of SilverStar at Feynman-diagram-gluon-radiation.svg, updated by Joel Holdsworth, CC BY 2.5,
https://commons.wikimedia.org/w/index.php?curid=1764161
Figure 2.1(l): By user: ComputerGeezer and Geof – Modified from :de:Bild:Venturirohr.jpg, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=8957957
Figure 2.1(r): By Thierry Dugnolle [CC0], from Wikimedia Commons
Figure 2.2(l): By Bob Cahalan, NASA GSFC [Public domain], via Wikimedia Commons
Figure 2.2(ra): By Astrobob – Own work, CC BY-SA 4.0,
https://commons.wikimedia.org/w/index.php?curid=44288890
Figure 2.2(rb): By NASA / JPL-Caltech / Space Science Institute [Public domain], via Wikimedia Commons
Figure 2.5: By Flanker – Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=3602085
Figure 2.5: Fair use, https://en.wikipedia.org/w/index.php?curid=13028554
Figure 2.7: By Source, Fair use,
https://en.wikipedia.org/w/index.php?curid=17357547
Figure 2.8(l): Zakharov AP and Pismen LM, 2018, Phys. Rev. E 97: 062705. Reproduced by permission, copyright by American Physical Society

L. Pismen, *The Swings of Science*,
https://doi.org/10.1007/978-3-319-99777-3

Figure 2.8(r): By Source, Fair use,
https://en.wikipedia.org/w/index.php?curid=23893955
Figure 2.10: By FF-UK – Own work, CC BY-SA 4.0,
https://commons.wikimedia.org/w/index.php?curid=59542523
Figure 3.1(r): By Prof. Jos van den Broek. CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=10970381
Figure 3.4(l): By Peteymills – Own work, CC0,
https://commons.wikimedia.org/w/index.php?curid=18966717
Figure 3.6(l): By Chetvorno – Own work, Public Domain,
https://commons.wikimedia.org/w/index.php?curid=11762440
Figure 3.6(r): By KCL.png: Pflododerivative work: M0tty (talk) – KCL.png, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=16594174
Figure 3.8(l): By Alksub at the English language Wikipedia, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=2756944
Figure 3.8(r): By Larryisgood – Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=17639292
Figure 3.9: By Alexis Monnerot-Dumaine – Author, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=1162764
Figure 4.1: By Libb Thims (talk) – Own work, Public Domain,
https://commons.wikimedia.org/w/index.php?curid=11437332
Figure 4.2(inset): By User:Htkym – Own work, CC BY 2.5,
https://commons.wikimedia.org/w/index.php?curid=1625737
Figure 4.3: By Daderot at English Wikipedia – Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=507663
Figure 4.4: Illustration by V. Chernov (Pismen, 1973)
Figure 4.5: By Haltopub – Own work based on Benzene Structural diagram.svg and Ouroboros-simple.svg, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=29661966
Figure 4.6: Illustration by V. Chernov (Pismen, 1973)
Figure 4.7(r): By Thomas Shafee (Own work) CC BY-SA 4.0
https://creativecommons.org/licenses/by-sa/4.0, via Wikimedia Commons
Figure 5.1(l): By Psiĥedelisto, based on version by Dbuckingham42 – Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=41724357
Figure 5.1(r): By Wilson Bentley – Plate XIX of "Studies among the Snow Crystals ..." by Wilson Bentley, "The Snowflake Man." From Annual Summary of the "Monthly Weather Review" for 1902, Public Domain,
https://commons.wikimedia.org/w/index.php?curid=22130
Figure 5.2(l): By Inductiveload – Own work, Public Domain,
https://commons.wikimedia.org/w/index.php?curid=5839079
Figure 5.2(c): By Materialscientist – Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=10094837
Figure 5.3(l,c): By Diamond_and_graphite.jpg: User:Itubderivative work: Materialscientist (talk) – Diamond_and_graphite.jpgFile:Graphite-tn19a.jpg, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=7223557

Figure 5.3(r): By User Qwerter at Czech wikipedia: Qwerter. Transferred from cs.wikipedia to Commons by sevela.p. Translated to english by Michal Ma?as (User:snek01). Vectorized by Magasjukur2 – File:3D model hydrogen bonds in water.jpg, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=14929959
Figure 5.4 By Javier Bartolomé Vílchez (Author of original image: David Gabriel García Andrade) – Modification of an image subject to Commons, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=4450178
Figure 5.5(l): By Matthieumarechal, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=4623701
Figure 5.7(c): By PD (Wikipedia english) GFDL
http://www.gnu.org/copyleft/fdl.html, or CC-BY-SA-3.0
http://creativecommons.org/licenses/by-sa/3.0/, via Wikimedia Commons
Figure 5.8(la): By Major measure – Own work, Public Domain,
https://commons.wikimedia.org/w/index.php?curid=5393833
Figure 5.8(lb): By Emmanuelle rio slr – Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=17297500
Figure 5.8(r): By FlagSteward, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=3748654
Figure 5.9(l): By MesserWoland – own work created in Inkscape, based on the graphics by Daniel Stiefelmaier, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=1353236
Figure 5.9(r): By Na2jojon (Own work) [Public domain], via Wikimedia Commons
Figure 5.10 Weinstein O and Pismen LM, 2008, Math. Model. Nat. Phenom. 3: 98–107. Reproduced by permission, copyright by EDP Sciences
Figure 5.11: By Kebes – Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=4170779,
https://commons.wikimedia.org/w/index.php?curid=4170810,
https://commons.wikimedia.org/w/index.php?curid=4170873
Figure 5.12(l,c): CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=1568630
Figure 5.12(r): Courtesy Prof. Masao Doi (Doi, 2013)
Figure 5.15: Courtesy Prof. Masao Doi (Doi, 2013)
Figure 5.16(l): By Carl Henderson – Own work, CC0,
https://commons.wikimedia.org/w/index.php?curid=39441934
Figure 5.16(c): By MnMTD2k
https://commons.wikimedia.org/wiki/File:Reptation_theory.jpg, CC0,
https://commons.wikimedia.org/w/index.php?curid=39440696
Figure 5.16(r): By Hai2410 – Own work, Public Domain,
https://commons.wikimedia.org/w/index.php?curid=10127776
Figure 5.17(l) Zakharov AP and Pismen LM, 2017. Soft Matter, 13: 2886. Reproduced by permission, copyright by Royal Society of Chemistry
Figure 5.17(ra) Zakharov AP and Pismen LM, 2016. Phys. Rev. E 93: 022703. Reproduced by permission, copyright by American Physical Society

Figure 5.17(rb) Zakharov AP and Pismen LM, 2018. Soft Matter, 14: 676. Reproduced by permission, copyright by Royal Society of Chemistry
Figure 6.2(l): By Darth Kule – Own work, Public Domain,
https://commons.wikimedia.org/w/index.php?curid=10555337
Figure 6.2(ra): By Bhutajata – Own work, CC BY-SA 4.0,
https://commons.wikimedia.org/w/index.php?curid=44144928
Figure 6.3(r): By Dhatfield – Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=4279886
Figure 6.4(lb,r): Illustration by V. Chernov (Pismen, 1973)
Figure 6.5: Illustration by V. Chernov (Pismen, 1973)
Figure 6.6: By Nanite – Own work, CC0,
https://commons.wikimedia.org/w/index.php?curid=26707196
Figure 6.7 By User:Tatoute – Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=577575
Figure 6.8: By Adwaele at English Wikipedia, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=19838505
Figure 6.9(l): By Piotr Jaworski, PioM EN DE PL; POLAND/Poznań – Inspiration: Image:EXPULSION.png, Public Domain,
https://commons.wikimedia.org/w/index.php?curid=325237
Figure 6.9(r): By Mai-Linh Doan (self photo) GFDL
http://www.gnu.org/copyleft/fdl.html, CC-BY-SA-3.0
http://creativecommons.org/licenses/by-sa/3.0/ or CC BY-SA 2.5-2.0-1.0
https://creativecommons.org/licenses/by-sa/2.5-2.0-1.0, via Wikimedia Commons
Figure 6.10: By Tem5psu, CC BY-SA 4.0,
https://commons.wikimedia.org/w/index.php?curid=38870945
Figure 7.1(r): By Krauss – Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=32611460
Figure 7.3(l): by ESRL OSD, https://www.esrl.noaa.gov/psd/outreach/education/science/convection/Pattern_s2.html
Figure 7.3(r): by Antigng, Own work, the Wikimedia Commons.
Figure 7.4: Golovin et al (2002). Reproduced by permission, copyright by World Scientific.
Figure 7.5(l): By X4or. Vectorized by Magasjukur2
File:CouetteTaylorSystem.png, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=14860138
Figure 7.5(c): By MyName (Dhollm (talk)) – Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=11553612
Figure 7.5(ra): Dror H and Pismen LM 1994, Chem. Eng Sci 49: 1119–1129. Reproduced by permission, copyright by Elsevier
Figure 7.5(rb): By Jojo V – Own work, CC BY-SA 4.0,
https://commons.wikimedia.org/w/index.php?curid=63603579
Figure 7.6(c): By Huboedeker at English Wikipedia, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=39072751,
Figure 7.6(r): By Paul C. Foster – Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=17262314

Figure 7.8(l): By Simpsons contributor (talk) – Own work (Original text: I (Simpsons contributor (talk)) created this work entirely by myself), Public Domain, https://commons.wikimedia.org/w/index.php?curid=18284509
Figure 7.9(1): Courtesy of Prof. Ronald Imbihl
Figure 7.9(2,3,6): NASA image, Public Domain
Figure 7.9(4): By Crosslink – Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=15554354
Figure 7.9(7): Hugh Venables/Shingle ridges on Orfordness, via Wikimedia Commons
Figure 7.9(9–11) Golovin AA and Pismen LM, 2004, Chaos 14: 845–854. Reproduced by permission, copyright by American Institute of Physics
Figure 7.10(l): Hagberg et al (2006). Reproduced by permission, copyright by Elsevier
Figure 7.10(r): By Chiswick Chap – Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=19437044
Figure 7.11(la): By Paul from Enschede, The Netherlands – Dendritic Copper Crystals, CC BY-SA 2.0, https://commons.wikimedia.org/w/index.php?curid=18829701
Figure 7.11(lb): By Claire Trease, Kingston University (Own work) [CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0)], via Wikimedia Commons
Figure 7.11(ca): By Bert Hickman – http://www.capturedlightning.com, Attribution, https://commons.wikimedia.org/w/index.php?curid=205440
Figure 7.12: Simulation of the Rayleigh–Taylor instability by Shengtai Li and Hui Li (LANL), Public Domain.
Figure 8.3 by User:Avsa, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=642216, https://commons.wikimedia.org/w/index.php?curid=642220, https://commons.wikimedia.org/w/index.php?curid=642221
Figure 8.5(c): By Gato ocioso – Own work, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=3577575
Figure 8.10(l): By Richard Sear - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=65111129
Figure 8.10(c): Golovin et al (1995). Reproduced by permission, copyright by American Chemical Society
Figure 8.10(r): By Xammm – Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=38592571
Figure 8.11(l): By D. Dibenski – images.fws.gov ([1]), Public Domain, https://commons.wikimedia.org/w/index.php?curid=3440966
Figure 8.11(r): By Brocken Inaglory – Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=7181229
Figure 9.1(r): By C. Fukushima and J. Westerweel, Technical University of Delft, The Netherlands (Own work) CC BY 3.0 http://creativecommons.org/licenses/by/3.0, via Wikimedia Commons
Figure 9.3(l): By Georgepehli – Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=15331086

Figure 9.3(r): By Student0531 – Own work,
https://commons.wikimedia.org/wiki/File:Haripins.png
Figure 9.5(la): By Rpgch CC BY-SA 4.0
https://creativecommons.org/licenses/by-sa/4.0, from Wikimedia Commons
Figure 9.5(lb): By Boumphreyfr – Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=7099928
Figure 9.5(r): By OpenStax
https://cnx.org/contents/FPtK1zmh@8.25:fEI3C8Ot@10/Preface, CC BY 4.0,
https://commons.wikimedia.org/w/index.php?curid=30015035
Figure 9.6(la): By JWSchmidt at the English language Wikipedia, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=12469419
Figure 9.6(lb): By OpenStax College – Anatomy & Physiology, Connexions Web site http://cnx.org/content/col11496/1.6/, Jun 19, 2013, CC BY 3.0,
https://commons.wikimedia.org/w/index.php?curid=30131299
Figure 9.6(ra): By Original uploaded by Bulbeck (Transfered by Vojtech.dostal) – Original uploaded on en.wikipedia, CC BY 2.5,
https://commons.wikimedia.org/w/index.php?curid=13338712
Figure 9.6(rb): By Fbustos1 GFDL http://www.gnu.org/copyleft/fdl.html or CC BY-SA 4.0-3.0-2.5-2.0-1.0 https://creativecommons.org/licenses/by-sa/4.0-3.0-2.5-2.0-1.0, from Wikimedia Commons
Figure 9.8(l): By TimVickers at English Wikipedia – Transferred from en.wikipedia to Commons, Public Domain,
https://commons.wikimedia.org/w/index.php?curid=2319237
Figure 9.8(r): By LadyofHats Mariana Ruiz Villarreal – Own work. Image renamed from Image:Facilitated_diffusion_in_cell_membrane.svg, Public Domain,
https://commons.wikimedia.org/w/index.php?curid=3981034
Figure 9.9(ra): By Thomas Shafee – Own work, CC BY 4.0,
https://commons.wikimedia.org/w/index.php?curid=39423556
Figure 9.10: Courtesy TM/R Design
Figure 9.14: By M. Readey CC BY-SA 3.0
https://creativecommons.org/licenses/by-sa/3.0, from Wikimedia Commons
Figure 9.15(l): By user:Looie496 created file, US National Institutes of Health, Public Domain, https://commons.wikimedia.org/w/index.php?curid=8882110
Figure 9.15(ra): By MethoxyRoxy CC BY-SA 2.5
https://creativecommons.org/licenses/by-sa/2.5, from Wikimedia Commons
Figure 9.15(rb): adapted from Original by en:User:Chris 73, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=44114666 and BruceBlaus, Own work, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=29452220
Figure 9.16: CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=184098,
https://commons.wikimedia.org/w/index.php?curid=466264
Figure 10.1: Illustration by V. Chernov (Pismen, 1973)