2

Fundamental physics: elementary particles and processes

This chapter serves to familiarize the Student with the physics of elementary particles, where new concepts are introduced in their historical context and without a precise, technical definition. The subsequent chapters will clarify these concepts with more details, examples and applications.

2.1 The subject matter

The task of elementary particle physics is explaining of what and how the World is fundamentally made. Amazingly, and almost exactly in a Democritean sense, substance (tangible matter) comprises tiny particles, and our task includes a coherent classification:

1. both a systematic inventory of these “elementary particles,”
2. and an understanding of the “elementary processes” between them,
   i.e., their “fundamental interactions.”

In principle, these fundamental interactions determine how collections of otherwise independent elementary particles bind into ever larger structures, to macroscopical and even astronomical proportions. However, all except the teeniest in this hierarchy of structures are outside the scope of this subject.

One must *actively* keep in mind that the seemingly homogeneous and continuous substance consists of only a few types of particles, amongst which each one occupies but a tiny volume, and between which most of the space is practically empty. Less than a *trillionth* of the volume of any given substance is occupied by the particles forming the substance. Corresponding to their tininess, these particles come in fantastic numbers, and these countless copies are all identical. Not only “practically equal,” but two particles of the same kind really cannot possibly be distinguished from one another: any one of the 10^{29-30} electrons in our body is identical and exchangeable with any of the other electrons. It is absolutely impossible to distinguish one electron from another, except by the state in which an electron is and by its interactions with the rest of the considered system. The same holds for protons and neutrons.

We will see that elementary particles are determined by their types of interaction with other particles. The seeming void between particles is in fact filled with interaction *fields*: in ordinary substances, this is the electromagnetic field. In this sense, it is incorrect to focus exclusively on particles, however elementary. However, the whole image of particles is just a picturesque *model*, a caricature; we must recall the quantum nature of Nature, and the particle–wave duality. Elementary particles must be thought of as neither particles nor waves, but as objects that in certain circumstances appear as very tiny and precisely localized particles, and in other circumstances look like dispersed, continuous waves. Similarly, the electromagnetic field is an object that in certain circumstances behaves as a continuous and spatially very much extended wave, but in other circumstances it looks like a well-localized elementary particle, the photon. Thus, there is no conceptual difference between the objects that we most often encounter as either particles, waves or fields; in field theory, all objects are represented by fields, the quanta of which may be particles.

— ❦ —

We manipulate macroscopic objects directly. Thus, Coulomb could experiment with mica spheres suspended by silk thread, and Cavendish with lead spheres suspended on a torsion swing. Particle physicists, however, cannot catch an electron with tweezers or string up a few protons on a thread. Experiments in elementary particle physics are thus reduced to studying (**1**) collisions and scattering,[1] (**2**) decays, and (**3**) bound states of elementary particles. The laws and rules of interactions are then reconstructed from the results of such studies.

As is well known, if any relative speed of any two sub-systems is comparable with the speed of light in vacuum, one must use relativistic mechanics. Also, if the Hamilton action[2] for a given process is comparable with $\hbar$, one must use quantum mechanics. In our case, we have to use physics that is both relativistic and quantum, i.e., quantum field theory. On the other hand, in an introduction such as this, we do not have to introduce the whole mathematical–technical apparatus of field theory, but rely as much as possible on picturesque models and analysis that is *conceptually* not much more demanding than the usual mathematical apparatus of non-relativistic quantum mechanics.

Some of the characteristics of elementary particle physics are essentially relativistic effects, while other properties stem from quantumness. For example, the 4-vector of energy and momentum (henceforth, "*4-momentum*") is always conserved in so-called real states – i.e., states that may be observed and measured, but mass is not. On the other hand, in *virtual states* the 4-momentum conservation laws need not hold; see Section 2.4.2. Also, Nature's relativity permits the existence of particles with identically vanishing mass: The particle of the electromagnetic field, the photon, makes no sense in non-relativistic physics; see p. 94. Moreover, the combination of relativity and quantum theory leads to results that can be obtained in neither the theory of relativity nor quantum mechanics. The existence of antiparticles, the proof of Pauli's exclusion principle (1925) and the so-called spin-statistics theorem and the so-called "CPT-theorem" all stem from the *combination* of quantum and relativistic, and all relativistic quantum models must include them.

By about 1978, the so-called Standard Model had taken form in elementary particle physics; it encompasses all phenomena involving the elementary particles and their interactions as known to date, and is in full agreement with the experimental data observed in the last three decades; see [307, 221, 422, 159, 423, 538, 250, 243] and also [458] for a first-person account of the 1960/70s excitements from an experimentalist's vantage point. Our main goal is to acquaint ourselves with that Standard Model and the basic principles of its structure, such as the gauge

---

[1] While the terms "collision" and "scattering" are often used interchangeably, the former will here tend to refer to the physical event of colliding or its bringing about, while the latter will tend to refer to the process and its results, often focusing on the individual particles involved, and often being inelastic.

[2] This is indeed *Hamilton's principal function*, the time-integral of the Lagrangian, familiar to Students from classical physics, where the integrand determines the physical system, and the boundary data and limits of integration specify the process considered.

principle, which is at the foundation of all fundamental interactions and links symmetries[3] with conservation laws via Noether's theorem.

That is, here we are interested in the *theoretical* physics of elementary particles and their elementary processes, via fundamental interactions. However, we must first, even if briefly, turn to the experimental aspects – to know what it is that we have set out to explore and describe theoretically.

2.2 Elementary particles: detection and predisposition in experiments

2.2.1 Production

Most instrumentation used in experimental elementary particle physics is familiar from the literature in nuclear physics, so only a brief review is given here.

Producing electrons for laboratory use is almost trivial: Metals, when heated or irradiated with UV light, emit electrons, which are then easy to "catch" and direct with electric and magnetic fields. Protons – when needed, say, in a beam – may be produced by ionizing hydrogen. Since protons are charged, they can be directed with electric and magnetic fields, just as the electrons. On the other hand, since the electron mass is $\sim 1,836$ times smaller than the proton mass, the electrons may be neglected for many experimental purposes, so that a tank full of hydrogen is in practice a tank full of protons. Of course, many more particles were discovered in the past century, and these particles stem, mainly, from three sources:

Cosmic rays and their interaction with the atmosphere. It is not possible to identify the particular process at the source of any particular cosmic ray, nor do we know all the types of processes that create them; we do know, however, that particles of even very high energy incessantly bombard the Earth and collide with the atomic nuclei of atmospheric gases. Particles resulting from these collisions further collide with atomic nuclei of atmospheric gases, in cascading collisions. Clearly, this way of producing elementary particles is completely uncontrolled and subject to happenstance. However, it *is* a source of extremely high-energy particles, which we otherwise cannot produce in the lab. Besides, this source is also completely free.

Nuclear reactors and sources Atomic nuclei in radioactive materials spontaneously decay and in such processes not infrequently emit neutrons, α-particles (helium atomic nuclei), β-particles (electrons or positrons, depending on the source) and γ-particles (photons). Also, irradiating materials with so-called synchrotron (electromagnetic) radiation frequently either directly produces new particles, or makes those materials radioactive.

Particle accelerators and colliders The basic idea is to direct and accelerate previously produced particles along well-established paths, and then either bombard a target with these particles, or direct two such beams at one another.[4] Let us mention here Van de Graaff's machine, Cockcroft and Walton's linear accelerator, Lawrence's cyclotron, and finally Wildröe and Touschek's betatron. In this last device, oppositely charged particles are accelerated in opposite directions and within nearly identical and approximately circular paths. These then collide at the intersections of these paths; this is the basic idea in contemporary colliders. Since particles are directed and accelerated using electric and magnetic fields, this clearly applies only to charged particles. In contemporary practice, particles are mostly directed along circular paths – following huge circular tunnels of miles-long radii. Thus, particles that "missed" in one "turn" get to collide again in the next one. As we will show in Section 3.2.2, the energy

---

[3] Symmetries have the mathematical structure of groups, so the mathematical subject of group theory turns out to be very useful in studying and using this structure. Appendix A provides a telegraphical review of the most needed results from this mathematical subject.

[4] Apparati for colliding beams of particles against each other are called colliders.

available in such collisions is "used" for creating new particles. Creation of lighter particles of course requires less energy, so that lighter particles are produced and discovered more easily, and the heavier ones are harder – or have not yet been created/discovered.

2.2.2 Nomenclature

In 1899, after having become known for his "investigations into the disintegration of elements and the chemistry of radioactive substances" (for which he would be awarded the Nobel Prize in Chemistry in 1908), Ernest Rutherford classified the radioactivity emitted by natural samples as α- and β-rays, distinguishing them by their penetrating power. Four years later, he found that radium, discovered in 1900 by Paul Villard, emitted a type of radiation that surpassed both α- and β-rays in penetrating power and named it γ-rays. By this time, he had (**1**) explaned that radioactivity in natural samples is caused by spontaneous disintegration of the atoms of the sample, (**2**) identified the exponential decay law and its application to use the constant decay rate as a "clock," and (**3**) identified the particles of α-rays as probably fully ionized helium atoms, i.e., helium nuclei. His classification was merely refined over the years and is still in use:

α-rays and particles are helium nuclei and consist of two protons and two neutrons.

β-rays and particles may be either electrons (also known as *cathode rays*), or positrons, depending on the process that created them. For example, the negatively charged cathode has a surplus of electrons, which a strong electric field may be able to free from the cathode and direct as a cathode ray. In nuclear processes, the so-called weak interaction can produce both electrons (in the β-decay of the neutron) and positrons (in the β-decay of the proton[5]), which are then emitted from the source material.

Table 2.1 The names of various bands of electromagnetic radiation

Name	Frequencies	Wavelength	Energy
γ-rays	> 30 EHz	< 10 pm	> 124 keV
Hard X-rays	3 – 30 EHz	10 – 100 pm	12.4 – 124 keV
Medium X-rays	0.3 – 3 EHz	0.1 – 1 nm	1.24 – 12.4 keV
Soft X-rays	30 – 300 PHz	1 – 10 nm	0.124 – 1.24 keV
Ultraviolet rays	0.79 – 30 PHz	10 – 380 nm	3.27 – 124 eV
Visible light	400 – 790 THz	380 – 750 nm	1.65 – 3.27 eV
Near infrared	30 – 400 THz	0.75 – 10 μm	0.124 – 1.65 meV
Medium infrared	3 – 30 THz	10 – 100 μm	12.4 – 124 meV
Far infrared	0.3 – 3 THz	0.1 – 1 mm	1.24 – 12.4 meV
Radio and micro-waves	< 0.3 THz	> 1 mm	< 1.24 meV

Standard prefixes: $\mathrm{E} = 10^{18}$, $\mathrm{P} = 10^{15}$, $\mathrm{T} = 10^{12}$, $\mathrm{k} = 10^{3}$, $\mathrm{m} = 10^{-3}$, $\mu = 10^{-6}$, $\mathrm{n} = 10^{-9}$, $\mathrm{p} = 10^{-12}$.

γ-rays and particles (photons) are the known quanta of electromagnetic radiation. More precisely, Table 2.1 shows the division of the electromagnetic radiation spectrum and names the various bands. Traditionally, "γ-rays" meant electromagnetic radiation that stems from spontaneous nuclear γ-decay, while "X- or Röntgen-rays" referred to electromagnetic radiation produced artificially. Initially, γ-rays were meant to denote electromagnetic radiation with energies higher than in X-rays, but contemporary accelerators produce X-rays with energies far above the energy of typical γ-rays, and these two bands overlap. By tradition and for simplicity,

---

[5] A free proton cannot decay into a positron and a neutron because of energy conservation. *[✎ Verify.]*

high-energy electromagnetic radiation is called γ-radiation regardless of its origin. In fact, "γ" is used as a symbol for the photon, the particle (quantum) of electromagnetic radiation in general and regardless of its frequency (energy) and wavelength.

Independently of this nomenclature, particles in relativistic physics are also classified depending on their mass and possible speeds of propagation, as compared with photons:

Tardion is a particle that in vacuum moves slower than light and has a real and positive mass; all known matter (and antimatter!) consists of tardions; see below.

Luxon is a particle that moves in vacuum at the speed of light in vacuum, c, and has no mass. All particles mediating gauge interactions that correspond to unbroken gauge symmetries are luxons.

Tachyon is a particle that moves in vacuum faster than light, and has an imaginary mass. The emergence of tachyons indicates that the ground state of the system (i.e., vacuum) is not stable [☞ Digression 7.1 on p. 261].

This classification refers to the Lorentz-invariant mass, defined precisely in Section 3.1.3.

2.2.3 Detection

Particle detection relies on the interaction between the particle and its environment, so that all detectors are more or less selective.

Geiger counter detects ionizing radiation, usually β- and γ-radiation, but there exist models that can also detect α-particles. It consists of a tube filled with an inert gas (usually helium, neon or argon, with halogen additives) that becomes conducting when radiation (partially) ionizes it. The spark created between the electrodes when the gas becomes conducting because of ionization provides a signal that is amplified by a cascading array of electrodes. The so-produced current is usually shown on a galvanometer, a pilot-lamp or by a speaker – hence the characteristic crackling. As the density of the gas in the tube is relatively small, very high-energy particles pass through without detectable interaction.

Neutrons are neutral and by themselves do not trigger the Geiger counter; however, they can be detected indirectly, by using boron trifluoride and a moderator that slows neutrons and in the process creates (charged and easily detectable) α-particles.

Scintillation counter consists of a transparent crystal and a fluorescent material that reacts to ionizing radiation; a sensitive photo-multiplier is used to amplify the signal.

Cherenkov counter is based on the fact that there exist materials through which some particles can travel faster than light, albeit slower than the speed of light in vacuum. Such particles also interact with this material and emit electromagnetic radiation, which then forms a cone-shaped shock wave with the opening angle $\theta_c = \arccos(c/nv)$, where v is the speed of the particles and n is the refraction index of this material. This effect was discovered by Pavel Cherenkov, while Ilja Frank and Igor Tamm explained it theoretically, for which the three of them shared the 1958 Nobel Prize.

Cloud chamber also known as the Wilson chamber, after Charles T. R. Wilson, is filled with super-cooled vapor (water or alcohol), in which the passing particle engenders condensation. This forms a sequence of condensed droplets that faithfully trace the particle's passage.

Bubble chamber (invented by Donald Glaser) is filled with superheated liquid (usually liquid hydrogen at $-253\,^{\circ}$C, propane or some other appropriate liquid), in which the passing particle engenders evaporation. As in the (Wilson) cloud chamber, this forms a sequence of evaporated bubbles that faithfully trace the particle's passage.

Spark chamber (invented by Shuji Fukui and Sigenori Miyamoto) is woven through with wires that are kept at various voltages, in very short (∼10 ns) impulses. When a particle passes between two wires and ionizes the gas between the wires, the gas (specially chosen for this purpose, usually an argon–methane mixture) becomes conducting and a spark jumps between the wires; the brevity of the voltage pulses prevents an avalanche of sparks. The spatial distribution and time sequence of the sparks faithfully trace the particle's passage and also depict its speed.

Proportional chamber is Georges Charpak's modification of the spark chamber that enables measurement of the quantity of ionized gas, which is proportional (whence the name) to the kinetic energy of the particle that produced the ionization. Charpak received the 1992 Nobel Prize for this invention.

Photographic emulsion contains molecules of silver halide, which react with passing charged particles. When the film is processed afterwards, the trace in the emulsion faithfully depicts the particles' paths. Clearly, this method captures faithfully only two-dimensional events that happen to be coplanar with the emulsion.

2.2.4 Predisposition in experiments

Experiments in elementary particle physics grew from the relatively humble confines of individuals' labs, such as Rutherford's in Manchester at the turn of the twentieth century, into humongous multi-national installations such as CERN. This evolution of experimental physics has side-effects that in many ways limit the scope, variety and intellectual freedom of the experiments and so limit the advancement of elementary particle physics, and even physics in general.

Even in principle, experiments are performed to test one concrete hypothesis or another, and hypotheses are of course limited by the imagination of the physicists who design the experiments. This creates a selection effect in experimental science: On one hand, we can derive/compute a result from a given concrete theoretical model and then design an experiment to test this result. Or, we can re-test some earlier result, but to a greater precision than was possible up to that point. Not infrequently, the crucial improvements in such experiments require an ingenuity and inspiration that is rightfully impressive and even awesome. However, the fact remains that such experiments – for the most part – "only" test existing/known theory.

On the other hand, experiments may also be designed to test hypotheses that are inspired by science fantasy or even child's play, which start with "what if . . . " questions, where the ellipses represents a hypothesis not limited by any concrete result from any concrete theoretical model. One would expect *such* experiments to have a much larger chance of discovering wholly unexpected effects and phenomena. To list but a few examples:

1. Thales (of Miletus, *c*. 620–625 BC!) noted that, after being rubbed against fur, amber attracts particles of dust.
2. The more systematic experiments of Alessandro Volta (1745–1827) showed that frog legs were induced to twitch by poking them with wires of *certain* metals, although the legs were cut off from the frogs and so were evidently dead.
3. Hans Christian Ørsted (1777–1851) observed that the magnetic compass needle changes its direction when brought near a wire through which a current is passing.
4. Henri Becquerel (1852–1908) discovered that certain (radioactive) materials affect photographic material via means wholly invisible to the human eye.

Many discoveries have occurred exactly through such "unbridled" and even entirely *accidental* activity. Had it not been for such spontaneously freewheeling experiments, who knows if electrodynamics would ever have developed into the theoretical model (and its ubiquitous practical applications!) that we know today? It is certain, however, that the history of physics would have turned out very differently.

Returning to elementary particle physics: evidently, such “unbridled” experiments were possible to design up to the mid-twentieth century, and there did indeed occur some completely accidental discoveries. However, in a situation where new discoveries – such as are expected of the LHC experiment at CERN – require budgets, logistics and political agreements of dozens of countries and over several years and even decades, for “unbridled” and “accidental” experimenting there is no chance. Even when these big experiments are very carefully and diligently designed, the socio-political climate can cancel them – or worse, can nix them half-way through (as happened with the Superconducting Super Collider, in 1993) – because of reasons that are not directly related to physics and research.

> ***Conclusion 2.1*** *This very nonlinear feedback between experimental physics, finance and the socio-political climate amplifies the selection effect, and produces an ever stronger and limiting influence – a predisposition – on the types of experiments that we can at least hope to perform. It is easily possible that this is one of the most important factors in the evident slow-down in experimental physics discoveries during the last quarter of the twentieth century and the first decade of the third millennium.*

As it is hard to believe that the financial and socio-political aspects will radically improve, experimental elementary particle physics must adapt if it is to survive; hopefully, by means of some radically new and clever (financially and socio-politically less limited) methodology [☞ also Digression 1.1 on p. 9].

2.3 A historical inventory of the fundamental ingredients of the World

The Democritean[6] idea of the smallest, indivisible constituents of the World was revived as the idea that there exists a smallest quantity of every chemical substance that retains the chemical properties unchanged. Carefully following the proportions in which chemical substances interact, chemists have established the existence of molecules and atoms, and even estimated their size, $\sim 10^{-10}$ m. The tininess of the molecules and atoms reciprocally implies their enormous number in macroscopically “normal” quantities of such substances. For example,

$$\text{there are } \approx 6.69 \times 10^{26} \text{ molecules in a (2 dl) glass of water.} \tag{2.1}$$

Molecules were known to consist of atoms, and – within the methods and techniques of nineteenth century chemistry – atoms are really indivisible: ἄτομος is ancient Greek for “uncuttable.”

2.3.1 The electron

In 1897, Joseph J. Thomson showed that atoms are not indivisible, and that a much smaller *electron* may be extracted from them: He showed that the cathode rays may be bent using electric and magnetic fields. The deflection in the electric field depends only on the electric charge of the particles that form the cathode ray, and the deflection in the magnetic field depends both on

---

[6] D. Griffiths [243] refers to Democritus as a *meta*physicist, and rightly so: many ancient Greek philosophers’ teachings have reached us as “armchair philosophy” and with no reference to experimental data in support. We thus carefully ascribe the *scientific* idea of atoms to the nineteenth century chemists, but the *inspiration* for that worldview to the *philosophy* of Democritus and Leucippus. It is worth noting that these Democritean ideas have spread mostly by way of the Latin epic *De rerum natura* [Titus Lucretius Carus, first century BC], which has been fully preserved till today. In six books, it presents the naturalist philosophy of the ancient Greek philosopher Epicurus, according to which the World consists of atoms that move in otherwise empty space.

the electric charge and on the particle speed. Thus, by manipulating the electric and the magnetic fields through which he let the cathode rays pass, Thomson determined that the ratio of the electric charge to the mass of the particles that form the cathode ray is several thousand times larger than the same ratio for any then known ion. It follows that cathode rays consist of particles of which either the electric charge is several thousand times larger, or the mass is several thousand times smaller than those of the then known ions.

Digression 2.1 If a particle enters with the speed v, in the direction of the positive x-axis, into a region with a constant electric field $\vec{E} = E_0\,\hat{e}_y$ and a constant magnetic field $\vec{B} = B_0\,\hat{e}_z$, it is affected by the Lorentz force

$$\vec{F} = q\,\vec{E} + q\,\vec{v} \times \vec{B} = q\,E_0\,\hat{e}_y + q\,v\,B_0\,(\hat{e}_x \times \hat{e}_z) = q\,\Big(E_0 - v\,B_0\Big)\,\hat{e}_y. \tag{2.2a}$$

If the particle does not deflect from its straight path, it follows that the total force vanishes, from which it follows that

$$v = \frac{E_0}{B_0}. \tag{2.2b}$$

If we now switch the electric field off, leaving the particle to follow a circular path of radius R, it follows that the magnetic (Lorentz) force provides the centripetal acceleration, so

$$q\,v\,B_0 = m\,\frac{v^2}{R} \qquad \Rightarrow \qquad \frac{q}{m} = \frac{v}{B_0\,R} = \frac{E_0}{B_0{}^2 R}. \tag{2.2c}$$

Between the two possible interpretations, the one stating that cathode rays consist of particles smaller than atoms seemed much more reasonable to Thomson. He referred to these particles as *corpuscles* and to their electric charge as the *electron*, but the latter name became universally accepted for the particles themselves.

Digression 2.2 Ironically, Walter Kaufmann (Berlin, Germany) had performed the same experiments as J. J. Thomson – at about the same time and more precisely! However, he did not leap to the same conclusion as Thomson. Adhering to the philosophical (epistemological) doctrine of *positivism*,[7] he could/would not conceive of the explanation in Democritean atomistic terms, explaining the cathode ray as a beam of particles too little to observe [553].

2.3.2 The proton

Since atoms were known to be electrically neutral, it followed that atoms consist of electrons and a positive part that is thousands of times more massive than electrons. The simplest supposition was that this positive part fills the volume of each atom and that the electrons are embedded within this positively charged mass; this was J. J. Thomson's so-called plum pudding model.

---

[7] Roughly, positivism restricts science to only address directly observable phenomena and discuss them only in terms of directly observable quantities.

In contrast, Ernest Rutherford[8] had, with his students Hans Geiger and Ernest Marsden, performed an epoch-making experiment in 1909: Bombarding a thin golden foil with α-particles, it was shown that atoms of gold (and so also of all other elements) are mostly empty. In 1911, Rutherford derived his classical formula, which is easy to rewrite into a contemporary form:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{e^2/4\pi\epsilon_0}{2m_\alpha v_0^2}\right)^2 \frac{1}{\sin^4(\theta/2)} \xrightarrow{\frac{e^2}{4\pi\epsilon_0}\to\alpha_e\hbar c} \left(\frac{\alpha_e\hbar c}{2m_\alpha v_0^2}\right)^2 \frac{1}{\sin^4(\theta/2)}, \tag{2.3}$$

where $\alpha_e = \dfrac{e^2}{(4\pi\epsilon_0)\hbar c}$ is the fine structure constant, m_α the α-particle mass, v_0 their speed of approach to the foil, and θ the angle of deflection from the original direction of motion. The ratio $\dfrac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$ is called the differential cross-section, and gives the probability distribution as a function of the probe's deflection angle θ. Besides, α-particles are positively charged, and owing to the Coulomb repulsion can approach the positively electrically charged portion of the atom only to a distance b, which is determined from equating the energies:

$$\tfrac{1}{2}m_\alpha v^2 = \frac{1}{4\pi\epsilon_0}\frac{q_{\mathrm{Au}}\,q_\alpha}{b} \qquad \Rightarrow \qquad b = \frac{1}{4\pi\epsilon_0}\frac{2q_{\mathrm{Au}}\,q_\alpha}{m_\alpha v^2}, \tag{2.4}$$

where q_{Au} and q_α are the electric charges of the positive part of the gold atom and the α-particles. Direct measurements show that the minimal value for b was around 2.7×10^{-14} m, which is three to four orders of magnitude smaller than the size of the gold atom.

These initially established characteristics of the α-particle scattering pattern clearly "mapped" the Coulomb repulsive force field of the positively charged parts of the gold atom. Further experiments and more detailed analysis of this α-particle scattering pattern during the next decade managed to obtain indications of non-Coulomb scattering and so establish that the positively charged part of the gold atom is localized within a radius of only about 3.5×10^{-15} m.

This gave rise to the so-called planetary model of the atom, in which all atoms have a positively charged nucleus, of a radius $\sim 10^{-15}$ m, around which electrons revolve in orbits of radii $\sim 10^{-10}$ m. The nucleus of the simplest atom, hydrogen, was named *proton* by Rutherford.

Having obtained his PhD in May 1911 and spent six months at Cambridge working with J. J. Thomson, Niels Bohr came to the University of Manchester in May 1912 to work with Rutherford. By 1913, he had postulated that the electron's angular momentum in these orbits is limited to integral multiples of a constant, $\hbar$. With this ad hoc quantization, Bohr successfully computed not only the binding energies of the electron in the hydrogen atom, but also derived a general formula for the wavelengths of the photons emitted in transitions, and which was in full agreement with the observed series in the line spectra named after Balmer, Lyman, Paschen, etc.

Digression 2.3 In the last decades of the twentieth century it became clear that the integrality of the angular momentum in atoms of the hydrogen type – Bohr's ad hoc postulate – is also the essential characteristic that stabilizes atomic orbitals, and in fact a concrete manifestation of a general principle. The modern understanding of theoretical physics uses the language of geometry and topology; topological invariants are, roughly,

---

[8] Rutherford was J. J. Thomson's graduate student at Cambridge, but these atomic nuclei experiments were designed and conducted (a decade after Thomson's) at the University of Manchester, where Rutherford was chair of physics from 1907 – after eight years at McGill University in Canada during which he worked on radioactive materials and for which he was awarded the 1908 Nobel Prize in Chemistry. A decade later, in 1919, Rutherford succeeded Thomson as the director of the Cavendish Laboratory at Cambridge.

quantities that depend on integers in a critical way, so that invariance follows from the impossibility of a continuous change of these integral characteristic quantities.

In retrospect then, with the benefit of hindsight, we may conclude that Bohr's postulate must be accepted because it stabilizes the atomic orbitals. On the other hand, it may in turn be explained by (reduced to) the de Broglie particle–wave duality. *[✐ Why?]*

By 1917, Rutherford's experiments had shown that atoms can split in collisions, and by 1932, Rutherford and his students John Cockroft and Ernest Walton had developed experimental techniques to split some atoms in a fully controlled fashion.

The "natural" assumption that atomic nuclei then consist of just the right number of protons to neutralize the total charge of the electrons that are in orbit around that particular nucleus, however, did not find support in experimental data: Even in the late nineteenth century, it was known that the next atom by mass, the helium atom, has two electrons, but a mass that is not twice but *fourfold* larger than the hydrogen atom. Lithium is the next element, with three electrons, but the mass of its atom is *six* or *seven* times bigger than that of the hydrogen atom. This exemplifies the tendency of atoms to be two or more times heavier than the product of their atomic number in the periodic table and the mass of the hydrogen atom.

2.3.3 The neutron

The disproportionally larger masses of atomic nuclei were explained in 1932, when James Chadwick (Rutherford's student) discovered that atomic nuclei contain another type of particle, besides protons. As that other kind of particle is neutral, he called them *neutrons*. Being the building blocks of nuclei, protons and neutrons are collectively called *nucleons*.

Also, the existence of neutrons helped explain the mismatch in many heavier nuclei: it was known that the nuclear masses are generally (for the first few rows in the periodic table) about twice as large as the electric charge of the nucleus. This induced the supposition that nuclei are composed of twice as many protons as necessary to cancel the charge of the orbiting electrons, and that the surplus of positive charge in the nuclei is neutralized by additional electrons that are confined to the nucleus. However, it was also known that protons and electrons have an intrinsic angular momentum (spin) of $\frac{1}{2}\hbar$. In many nuclei the total sum of angular momenta from both the protons and the electrons – both those in the $\sim 10^{-10}$ m orbits and those (hypothetical ones) inside the nuclei – did not agree with experimental data on angular momenta.

For example, nitrogen-14 would in this model have 14 protons and 7 electrons in the nucleus, and 7 electrons in orbit, a total of 28 spin-$\frac{1}{2}$ particles. The total angular momentum of nitrogen-14 would then have to be an integral multiple of $\hbar$. By contrast, the measured value of the total angular momentum of nitrogen-14 atoms is always a half-integral multiple of $\hbar$. The discovery of the neutron made it clear that the nitrogen-14 nucleus consists of 7 protons and 7 neutrons, which guarantees the total angular momentum of the nucleus to be an integral multiple of $\hbar$ (and half-integral for the whole atom), in agreement with all observations.

Conclusion 2.2 *In hindsight, the 1932 theory of elementary particles now seems fantastically simple: The World consisted of* ***electrons****,* ***protons*** *and* ***neutrons****, it was already known that the first two of these interacted by exchanging* ***photons****, and it "only" remained to figure out how these particles form the bigger structures: atoms, molecules, etc. Of course, that turned out to be a very naive point of view, and as we will soon see, one that is rather far from the full picture.*

2.3.4 The photon

Newton's original idea that light consists of *corpuscles* did not stand the test of experiments of his day: By the twentieth century, the wave nature of light was fully obvious. This implied the same wave-like nature for all the different types of electromagnetic radiation, in the spectrum of which the visible light occupies but a tiny region.

True to our leitmotif from the introductory Section 1.1.1, the twentieth century began with a systematic demolition of this image: At the turn of the nineteenth into the twentieth century, Max Planck studied the problem of the so-called ultraviolet catastrophe. Classical statistical physics had up to that point in time perfectly explained all known thermodynamic processes. However, its application to electromagnetic waves emitted by all hot objects, such as a piece of ember or a star, produced nonsense – that the total power of the emitted radiation is *infinite*, as well as that the radiation intensity grows unboundedly with growing frequency. Experimental data with no exception show that the intensity does grow with frequency, but only up to a particular maximum (proportional to the temperature of the hot object) and decays thereafter. Also, it is patently obvious that the total power emitted by a hot object must be finite.

In 1900, Planck showed that the ultraviolet catastrophe may be avoided and the experimentally known spectra may be explained theoretically – if we assume that hot objects emit radiation in "packets," the *action* (roughly, energy times the duration of time) of which is an integral multiple of a constant, h – soon enough called the Planck constant. No other viable resolution of this problem has been found since.

In 1905, Albert Einstein showed that the photoelectric effect unambiguously indicates that electrons also absorb electromagnetic radiation in the same "packets," with *identically(!)* the same constant h. From this, Einstein concluded that not only is electromagnetic radiation both emitted and absorbed in such "packets," but that it also *exists* only in the form of such packets. This revolutionary idea encountered enormous resistance: Even a decade later and its quantal emission and absorption notwithstanding, most physicists agreed that electromagnetic radiation was nevertheless of wave-like nature.

In 1910, Peter Debye proved that Planck's result follows from supposing that the Fourier modes of the electromagnetic field have energies that are integral multiples of the $h\nu$ product. Nevertheless, only 15 years after that (in 1925, by which time the quantum nature of the electromagnetic radiation was largely accepted) had Max Born, Werner Heisenberg and Pascual Jordan correctly interpreted the electromagnetic field Fourier mode of energy $nh\nu$ as n particles with energy $h\nu$ each.

In the interim, Arthur H. Compton showed in 1923 that Einstein's claim of the quantum nature of electromagnetic radiation is the *only* known one that successfully explains the scattering of visible light and "soft" X-rays on free electrons. Compton showed that the analysis of the scattering as a collision of particles of light with electrons – by using the 4-momentum conservation laws – gives the formula that today bears his name:

$$\lambda' = \lambda + \lambda_c(1 - \cos\theta), \qquad \lambda_c = \frac{h}{m\,c} = \frac{2\pi\hbar}{m\,c}, \tag{2.5}$$

where λ_c is the so-called Compton wavelength for a particle of mass m. Compton's original analysis was meant for electrons, $m = m_e$, but it is obvious that the formula applies to the scattering of electromagnetic radiation on any free charged particle; see Exercise 2.4.2. The classical wave analysis of the scattering of electromagnetic radiation from charged particles, the so-called Thomson scattering, predicts a change in the radiation wavelength for sufficient radiation intensity. However, when the intensity is too small, Thomson's effect vanishes. By contrast, Compton's effect gives the correct change in the wavelength of the scattered photon regardless of the radiation intensity, and even for a single photon of energy $E_\gamma = 2\pi\hbar\nu = 2\pi\hbar c/\lambda$.

The name *photon* itself was given to the particles of light by the chemist Gilbert Lewis, as late as 1926, together with his proposal that photons can be neither destroyed nor created; the details of Lewis's proposal proved not to be what Nature has to offer, but the name stuck. Niels Bohr, Hendrik Kramers and John Slater as late as 1924 tried to "save" the non-corpuscular nature of the electromagnetic field, by proposing the so-called "BKS model" [70], which required that:

1. energy and momentum are conserved only in the average, but not in processes where a single charged elementary particle absorbs or emits electromagnetic radiation;
2. causality should not hold in such elementary processes.

The BKS model was swiftly given a "decent burial" [406], but inspired Werner Heisenberg in 1925 to co-develop with Max Born and Pascual Jordan the so-called "matrix mechanics" [547]. Also, Einstein received the Nobel Prize only in 1921 "for his services to theoretical physics, and especially for his discovery of the law of the photoelectric effect" – and not for the quantal understanding of light. It thus took over a quarter of a century from Planck's original hypothesis for physicists to accept the quantum nature of light.

2.3.5 Duality and locality

In 1924, de Broglie defended his doctoral dissertation with the fundamental idea of the *duality* between particles and waves (for which he received the 1929 Nobel Prize), whereby *every* particle of momentum $\vec{p}$ may be represented by a wave of

$$\text{wavelength}\quad \lambda = \frac{h}{|\vec{p}|}, \qquad \text{i.e., wave-vector}\quad \vec{k} = \frac{h}{\vec{p}^{\,2}}\vec{p}, \tag{2.6}$$

and vice versa. Thus, by about 1926, accepting the particle-like nature of light no longer implied abandoning the wave-like nature, but accepting a more general view whereby the objects we call electron, proton, neutron, photon, etc., under certain circumstances behave as particles, but as waves under other circumstances. We will see later, combining quantum theory and the special theory of relativity in field theory, all these objects may be unambiguously represented by appropriate *fields*, of which "particles" and "waves" are certain limiting forms.

— ❦ —

The quantum nature of electromagnetic radiation was at first very hard to accept, and even then mostly owing to the practical ease in explaining Compton's effect. It turns out, however, that such an understanding of electromagnetic radiation has a very deep consequence regarding the essential understanding of the fundamental electromagnetic interaction – and following this template, later, also the other interactions.

The classical understanding of the interaction between two charged bodies relies on the idea that each charged particle creates around itself an electric field. Then, on any other charged particle probing this field there acts a force equal to the product of the probing charge and the electric field being probed. This electric field then simply equals the Coulomb force per unit probing charge. By its definition, the electric field is thus a crutch for predicting the Coulomb force – which then *acts at a distance*. The role of the magnetic field is, in this classical understanding of the electromagnetic interaction, conceptually identical, and their sum gives the full Lorentz force.

One might argue that this action at a distance does not literally contradict the special theory of relativity, according to which all information propagates locally, from a point to another in a continuous space and at most at the speed of light in vacuum. That is, the establishing and all changes in this classical electromagnetic field travel at the speed of light in vacuum or slower than that when traversing a substance that interacts with the electromagnetic field. However, once established, the Coulomb field *instantaneously* produces the force upon a test charge, regardless of how far the test charge is from the source of the Coulomb field. It is this assumed instantaneity that does contradict the fundamental idea of locality, which is in turn woven into the special theory of relativity. Thus, *classical* field theory in fact does incorporate a conceptual contradiction.

The quantum nature of light replaces the concept of an everywhere present electromagnetic field (that instantaneously produces a force at a distance) with the concept wherein charged particles constantly emit and re-absorb photons, and this continual exchange of photons between two charged particles mediates the electromagnetic interaction between the two charges. The elementary interaction here is a charged particle emitting or absorbing a photon. Then, (**1**) a charged particle at some point emits a photon, (**2**) which travels at the speed of light in vacuum to another charged particle, (**3**) which then absorbs it. This photon has thereby mediated the interaction between the two charged particles. As this mediating photon cannot possibly be observed without changing the interaction it mediates, it is a *virtual particle*. The 4-momentum conservation laws cannot be applied to it, since neither energy nor momentum can be observed and measured so as to check. Therefore, there may well be an infinite number of emitted and absorbed virtual mediating photons, and only their combined effect is observed as the effective interaction between the two charged particles. The electromagnetic interaction between two charged particles thus occurs not via an instantaneous action at a distance, but at the speed of light and via the exchange of photons that *mediate* the electromagnetic interaction.

However, it should be clear that this is not a simple kinematic exchange. For example, two ice skaters (to limit the relevance of friction) throwing snowballs at each other certainly exert an interaction mediated by the snowballs. However, this purely kinematic method cannot describe an attractive force: electromagnetic forces may well also be attractive, and even more complicated than that when the charges move with respect to each other.[9]

It will turn out that understanding interactions as *mediated* is generally applicable and in fact a fundamental idea, which will lead to a unified understanding of all fundamental interactions [☞ Chapters 5–7 and 9 for details].

With this in mind, and for ease of locution, we will speak of elementary *particles*, but implicitly understand the de Broglie duality with waves, as well as the essentially more fundamental but technically more demanding representation by fields, the quanta (smallest packets) of which are the elementary particles.

2.3.6 Mesons

Elementary particle physics as described so far had no answer to the obvious question: What keeps the nucleons (protons and neutrons) within the atomic nucleus? It is clear that this force cannot be of electromagnetic origin – neutrons are neutral, and protons repel each other. What's more, this *strong nuclear force* must be stronger than the electromagnetic one, so as to overpower the Coulomb repulsion between protons – at least as long as they are within the atomic nucleus, i.e., at distances no larger than $\sim 10^{-15}$ m. However, at distances much larger than $\sim 10^{-15}$ m, this nuclear force must become negligible even between neutrons, and certainly as compared with the Coulomb repulsion between protons. Since the Coulomb force decays uniformly as $\sim 1/r^2$, it follows that at distances larger than $\sim 10^{-15}$ the strong nuclear force must decay suddenly, much faster than $\sim 1/r^2$.

In 1932, Werner Heisenberg proposed the formalism of *isospin*,[10] to explain the significant similarity between protons and neutrons: their masses differ only by 0.14%, they both have spin $\frac{1}{2}$,

---

[9] For example, the force with which a small magnetic dipole acts upon an approaching charged particle is orthogonal to the direction between the dipole and the charged particle in motion.

[10] The name itself was bestowed by Eugene Wigner, in 1937.

and the strong interaction within the atomic nuclei does not differentiate between protons and neutrons. Heisenberg thus assigned [☞ Section 4.3 for details]

$$p^+ \mapsto |\tfrac{1}{2},+\tfrac{1}{2}\rangle \qquad \text{and} \qquad n^0 \mapsto |\tfrac{1}{2},-\tfrac{1}{2}\rangle, \tag{2.7}$$

with formal operators $\vec{I}^2, I_3$ defined after those of angular momentum [☞ Appendix A], so that

$$\vec{I}^{\,2}|\tfrac{1}{2},\pm\tfrac{1}{2}\rangle = \tfrac{1}{2}(\tfrac{1}{2}{+}1)\,|\tfrac{1}{2},\pm\tfrac{1}{2}\rangle \qquad \text{and} \qquad I_3|\tfrac{1}{2},\pm\tfrac{1}{2}\rangle = \pm\tfrac{1}{2}\,|\tfrac{1}{2},\pm\tfrac{1}{2}\rangle. \tag{2.8}$$

This model implies that the *exchange* of isospin occurs by exchanging the isospin states

$$\begin{aligned} |1,+1\rangle : &\qquad |1,+1\rangle|\tfrac{1}{2},-\tfrac{1}{2}\rangle \to |\tfrac{1}{2},+\tfrac{1}{2}\rangle, \\ |1,0\rangle : &\qquad |1,0\rangle|\tfrac{1}{2},\pm\tfrac{1}{2}\rangle \to |\tfrac{1}{2},\pm\tfrac{1}{2}\rangle, \\ |1,-1\rangle : &\qquad |1,-1\rangle|\tfrac{1}{2},+\tfrac{1}{2}\rangle \to |\tfrac{1}{2},-\tfrac{1}{2}\rangle, \end{aligned} \tag{2.9}$$

so that by absorbing the state $|1,+1\rangle$, the state $|\tfrac{1}{2},-\tfrac{1}{2}\rangle = |n^0\rangle$ becomes $|\tfrac{1}{2},+\tfrac{1}{2}\rangle = |p^+\rangle$, etc. Overall, in all such processes isospin is conserved.

In 1934, Hideki Yukawa proposed the potential

$$V(r) = -g^2\,\frac{e^{-r/r_Y}}{r}, \tag{2.10}$$

where the coefficient g^2 is analogous to the product of two electric charges in Coulomb's law, and r_Y is the effective range of the force, for which experiments produce a value of about $r_Y \sim 10^{-15}$ m. Unlike the Coulomb potential that is established by photons – particles with no mass,[11] Yukawa's potential is mediated by particles with mass

$$m_\pi \sim \frac{\hbar}{r_Y c}, \qquad \text{so that} \qquad m_\pi \sim 150\text{–}200\,\text{MeV}/c^2, \tag{2.11}$$

and this mass is responsible for the exponential decay of the potential (2.10). Yukawa's proposal thus predicted a new particle, the *pion*, with a mass (2.11) that is between the electron mass ($m_e \approx 0.511\,\text{MeV}/c^2$) and the proton mass ($m_p \approx 938\,\text{MeV}/c^2$). Combining with Heisenberg's isospin proposal, Yukawa's theory predicts three pions:

$$\pi^+ \leftrightarrow |1,+1\rangle, \quad \pi^0 \leftrightarrow |1,0\rangle, \quad \pi^- \leftrightarrow |1,-1\rangle, \tag{2.12}$$

so that the relations (2.9) correspond to the processes

$$\begin{aligned} |1,+1\rangle|\tfrac{1}{2},-\tfrac{1}{2}\rangle \to |\tfrac{1}{2},+\tfrac{1}{2}\rangle, &\qquad \pi^+ + n^0 \to p^+, \\ |1,0\rangle|\tfrac{1}{2},\pm\tfrac{1}{2}\rangle \to |\tfrac{1}{2},\pm\tfrac{1}{2}\rangle, &\qquad \pi^0 + (p^+ \text{ or } n^0) \to (p^+ \text{ or } n^0), \\ |1,-1\rangle|\tfrac{1}{2},+\tfrac{1}{2}\rangle \to |\tfrac{1}{2},-\tfrac{1}{2}\rangle, &\qquad \pi^- + p^+ \to n^0. \end{aligned} \tag{2.13}$$

What is unusual as compared with electromagnetic interactions: unlike the photon, which is single and electrically neutral, Yukawa's proposal included *three* pions, with charges $-1, 0, +1$, and they would span the 3-dimensional (nontrivial!) representation of the non-abelian (non-commutative) isospin group $SU(2)_I$. This proposal is the forerunner of so-called non-abelian theories of gauge symmetry, which would be introduced two decades later by Chen-Ning Yang and Robert L. Mills, and independently also by Ronald Shaw in his PhD dissertation under Abdus Salam [473].[12]

---

[11] For a particle of mass m, we have that $E^2 = \vec{p}^{\,2}c^2 + m^2c^4$; for photons $m_\gamma = 0$, so that $E_\gamma = |\vec{p}_\gamma|c$.

[12] Ironically, Ernst Stückelberg had independently come up with a proposal very similar to Yukawa's, but Pauli's critique discouraged him from developing his idea. On the other hand, Pauli himself worked independently on a non-abelian generalization of electromagnetism, but was discouraged by his own critical views regarding difficulties in applications to weak interactions – which were resolved only much later by means of the Higgs mechanism – so that he never published on this topic [538].

In 1937, two groups of researchers (C. D. Anderson and S. H. Neddermeyer on the West Coast and J. C. Street and E. C. Stevenson on the East Coast of the USA) independently verified the existence of particles with a mass of the order of magnitude (2.11) when analyzing cosmic ray processes. Later and more precise measurement showed that the particles observed in cosmic rays have a mass very close to 100 MeV/c^2 – which is less than Yukawa's result (2.11), and soon enough different measurements started showing differing results. World War II interrupted these studies, but in 1946, it was shown in Rome, Italy, that the particles of mass $\sim$ 106 MeV/c^2 discovered in cosmic rays interact very weakly with atomic nuclei – completely contrary to the particles in Yukawa's proposal that were supposed to be the very mediators of the strong nuclear force!

In 1947, Robert Marshak and Hans Bethe proposed, and Cecil Powell (in collaboration with C. M. G. Lattes, H. Muirhead and G. P. Ochialini) also verified experimentally that cosmic rays actually involved two types of particles:

1. $\mu^{\pm}$-particles with a mass of $m_\mu \approx 106$ MeV/c^2, which interact very weakly with atomic nuclei and behave like a $\sim$206 times more massive copy of the electron;
2. π-particles with masses of 135 MeV/c^2 (π^0) and 140 MeV/c^2 ($\pi^{\pm}$), which interact strongly with atomic nuclei, and do fit Yukawa's prediction of 12 years earlier.

Later measurements and the quark model would show that the pions cannot be identified with the mediators of the strong interaction, although they do interact by means of the strong interaction, both amongst each other, and also with protons and neutrons; see the discussion leading to equation (6.77).

2.3.7 Antiparticles

Non-relativistic quantum mechanics was completed in only three years, 1923–6: it was conceptually clear that the Schrödinger equation could be adapted to every quantum system, whereupon it "only" remained to solve the differential equation subject to appropriate boundary conditions. Relativistic quantum theory, however, was a tougher nut to crack.

In 1927, Paul Dirac discovered the equation that bears his name, and which was supposed to describe free electrons. However, that differential equation has a solution of energy $E = -\sqrt{\vec{p}^{\,2}c^2 + m^2c^4}$ for every solution of energy $E = \sqrt{\vec{p}^{\,2}c^2 + m^2c^4}$. To avoid the preposterous possibility that the electron interminably loses energy as it successively falls into states of ever lower *negative energy*, Dirac initially proposed that all (infinitely many!) negative-energy states are filled, so that Pauli's principle prevents free electrons from falling into any of those filled states. This "sea" of infinitely many negative-energy electrons (one in each negative-energy state) is totally uniform, so that only the individual electrons with positive energy may be observed.

Furthermore, if any one of these negative-energy electrons were to acquire enough energy to become a positive-energy electron, in its former place in the infinite "Dirac sea" there would remain a "hole" with positive electric charge. Dirac initially hoped that these positively charged holes might be identified with protons, but Hermann Weyl soon showed that the inertial mass of these "holes" is equal to the mass of the free electrons, and so about 1,836 times too small for protons. The other problem with Dirac's idea of an infinite "sea" of electrons with arbitrarily negative energies is that the universe would have an infinitely negative total electric charge, and an infinitely large total mass. Ernst Stückelberg, and then Richard Feynman a little later and in more detail, re-interpreted the theory by introducing the concept of antiparticles (although Feynman attributes this invention to Dirac [166]): according to this by now standard understanding, an antiparticle is a particle traveling backwards in time (both with positive eneries), so that there is no infinitely deep and infinitely charged (for charged particles) Dirac sea [☞ Chapter 3.3].

Digression 2.4 The following is a simplified discussion from Ref. [166]: In all versions of the Fourier transformation there is a factor $e^{-i(E/\hbar)t}$. Changing the sign of energy E is here equivalent to flipping the direction of the passage of time t. Also, the Fourier transform $\widetilde{f}(\omega) := \frac{1}{\sqrt{2\pi}}\int \mathrm{d}t\, e^{i\omega t} f(t)$ has the key property that if $\widetilde{f}(\omega) \neq 0$ for only $\omega := (E/\hbar) \geqslant 0$, then $f(t)$ does not vanish in any continuous interval of time. It then follows that for a process that happens as a successive occurrence of two sub-processes (which are not observed/measured separately and independently but where all exchanged particles have non-negative energy), the time sequence of the two sub-processes depends on the (relativistic!) choice of observer's coordinate system and is not Lorentz-invariant. In turn, it then follows that:

1. antiparticles with positive energy are ordinary particles with positive energy traveling backwards in time;
2. the operations of *parity* (P), time-reversal (T) and charge conjugation (C) satisfy the relation $\mathsf{PT} = \mathsf{C}$ [☞ Section 4.2.3];
3. a particle–antiparticle pair may be created from vacuum and may (re-)annihilate;
4. for probability conservation:
   - (a) two fermions must not be in the same quantum state, so the creation and the re-annihilation of a fermion–antifermion pair contributes to the amplitude of probability negatively,
   - (b) two bosons can be in the same quantum state, and that increases the amplitude of probability of the (sub-)process.

See the diagrams (3.82) as well as Feynman's very intuitive and yet sufficiently detailed explanation in [166], where most of Feynman's "half" is dedicated to this connection.

In 1931, Anderson experimentally verified the existence of the *positron* and so also Dirac's theory. However, this implied that all other particles also must have their antiparticles,[13] of the same mass and of the opposite electric charge. Indeed, the antiproton was experimentally verified in 1955, and the antineutron the very next year.

Standard notation for antiparticles is the symbol of the particle with an over-bar: a proton is denoted by p, an antiproton by $\overline{p}$; n denotes a neutron, $\overline{n}$ an antineutron. However, charged leptons [☞ Table 2.3 on p. 67] are customarily denoted by means of the positive charge in the superscript: e^- denotes an electron while e^+ is a positron; μ^- denotes a muon while μ^+ is an anti-muon; τ^- is a tau lepton while τ^+ is the anti-tau lepton. Some of the neutral particles are their own antiparticles, such as the photon, γ, the neutral pion, π^0, and the weak nuclear interaction mediator, Z^0.

2.3.8 Crossing symmetry and detailed balance

There exists a general principle called ***crossing symmetry***, according to which for every process

$$A + B \to C + D, \tag{2.14a}$$

where A, B, C, D are particles partaking in the process, the following processes are also possible:

$$A \to \overline{B} + C + D, \tag{2.14b}$$

---

[13] Strictly, this prediction of antiparticles pertains only to particles of spin $\frac{1}{2}\hbar$, to which Dirac's equation applies.

$$A + \overline{C} \to \overline{B} + D, \tag{2.14c}$$

$$\overline{C} + \overline{D} \to \overline{A} + \overline{B}, \qquad \text{etc.}, \tag{2.14d}$$

provided that the ***kinematic*** (4-momentum) conservation laws permit them. In addition, the principle of ***detailed balance*** further predicts the existence of the reverse processes:

$$C + D \to A + B, \qquad \text{etc.}, \tag{2.14e}$$

again, provided that the kinematic conservation laws permit them. We will see later that the computations of the probability for these processes consist of two stages. The first is identical for all of the listed processes (2.14), while the second depends on the kinematics and the 4-momentum conservation laws and may well be drastically different.

> ***Comment 2.1*** *Note the difference between the "crossing symmetry" and the "principle of detailed balance" in their present use: The process (2.14d), obtained from (2.14a) using the crossing symmetry, certainly differs from the process (2.14e), which was obtained applying the principle of detailed balance. In this sense, the crossing symmetry permits "moving" any one particle amongst the results of a process into its antiparticle at the input of the process. On the other hand, the application of the principle of detailed balance is equivalent to reversing the direction of time: compare the process (2.14a) with (2.14e) [☞ Section 4.2.2].*

One of the direct applications of the crossing symmetry is the relationship between Compton scattering

$$e^- + \gamma \to e^- + \gamma \tag{2.15}$$

and electron–positron annihilation

$$e^- + e^+ \to 2\gamma, \tag{2.16}$$

which shows that the electron–positron annihilation produces *two* photons. That it cannot produce a single photon follows from 4-momentum conservation; see Exercise 3.2.7.

2.3.9 Neutrinos

Lisa Meitner and Otto Hahn had shown in 1911 that the β-decay of atomic nuclei seems to violate the energy conservation law.[14] In decays

$${}^{A}_{Z}X \xrightarrow{\beta} {}_{Z+1}^{\;\;A}Y + e^-, \qquad \text{such as} \qquad {}^{40}_{19}\text{K} \xrightarrow{\beta} {}^{40}_{20}\text{Ca} + e^-, \tag{2.17}$$

the electron's total relativistic energy (and also the magnitude of the linear momentum) is completely determined by the 4-momentum conservation law [☞ Digression 3.8 on p. 96]:

$$E_e = \Big(\frac{m_X^2 + m_e{}^2 - m_Y^2}{2\,m_X}\Big)c^2 \;=\; c\sqrt{m_e{}^2c^2 + \vec{p}_e{}^2}, \tag{2.18}$$

$$|\vec{p}_e| = c\,\frac{\sqrt{\big[(m_X+m_e)^2 - m_Y^2\big]\big[(m_X-m_e)^2 - m_Y^2\big]}}{2\,m_X}. \tag{2.19}$$

Measurements of all energies and linear momenta in β-decays of the type (2.17) showed that the value (2.18) is only the *maximal* value, and that the electron energy varies from case to case. Niels

---

[14] The neutron was first discovered in 1932, so that one did not know about its role in β-decays of atomic nuclei.

Bohr thus proposed that the 4-momentum conservation laws do not hold in processes involving such small particles.[15]

Opposing Bohr, Wolfgang Pauli proposed in 1930 that – so as to preserve the energy conservation law – β-decay (2.17) actually produces a third particle, with a very small mass and no electric charge, and so difficult to observe. However, this third particle carries away some of the energy, so that the measured values of the electron energy vary from case to case. Pauli proposed the name *neutron*, but was "scooped" for this name by Chadwick, who already used it (in published form by 1932) for the neutral particles that are a little heavier than the protons and which make up about half of the atomic nuclei of most elements.

In 1931, Enrico Fermi named Pauli's particle the *neutrino*[16] and by 1934 had published his theory of β-decay, based on the neutron decay

$$n \xrightarrow{\beta} p + e^- + \overline{\nu}_e, \tag{2.20}$$

which was very successful in describing the experimental observations and measurements. Only later was it established that Pauli's particle was in fact the antineutrino and of the electron species, as correctly denoted in the process (2.20).

2.3.10 Leptons

During 1937–47, particles of about 100–150 MeV/c^2 mass were found in photographs of cosmic ray induced processes, and these were at first identified with Yukawa's mediators of the strong interaction. However, in 1947, Cecil Powell showed that these photographs involve two very different kinds of particles: In a characteristic cascade decay (frequently found on the same photograph) we have

$$\begin{aligned} \pi^- \longrightarrow \mu^- &+ \overline{\nu}_\mu, && (2.21a)\\ \searrow\ & e^- + \overline{\nu}_e + \nu_\mu. && (2.21b) \end{aligned}$$

Using the derivation (3.44)–(3.49) below, the first decay (2.21a) shows that precisely one particle is not recorded in the photograph (because it is not charged), as the energy of the recorded muon is fixed. Again using the derivation (3.44)–(3.49) below, the second decay (2.21b) produces two invisible particles, because the energy of the visible electron varies. The invisible particles produced in the processes (2.21a) and (2.21b) have here been correctly denoted as $\overline{\nu}_\mu$ and $\overline{\nu}_e + \nu_\mu$, respectively, although in the analyses before 1962 [☞ Table 2.4 on p. 69] one did not know about the difference between the electron- and the muon-neutrino, nor which one in these processes is an *anti*neutrino.

In the first two decades of Pauli's proposal, theoretical proofs that neutrinos must exist abounded. However, no experimental verification was known.

In 1956, Frederick Reines and Clyde Cowan published the results of one of the first big "waiting experiments": a huge tank of water with detector-studded walls, where they waited to observe the so-called inverse β-decay, the process

$$\overline{\nu}_e + p^+ \longrightarrow n^0 + e^+, \tag{2.22}$$

guaranteed to exist by the crossing symmetry. By clever and detailed analysis of a large number of measurements, Reines and Cowan managed to provide an unambiguous experimental proof of the

---

[15] Supposedly [243], Bohr also vigorously fought Einstein's proposal that electromagnetic radiation exists in quanta; he opposed Dirac's electron theory and Pauli's neutrino proposal, ridiculed Yukawa's π-meson theory, and discouraged Feynman from his diagrammatic approach to quantum electrodynamics; he also advised Heisenberg against publishing his uncertainty relation [119].

[16] In Italian, "neutrino" is the diminutive of "neutron," i.e., the "little neutron."

neutrino's existence. Additionally, their analysis showed that antineutrinos interact with ordinary matter extraordinarily feebly: By contemporary estimates (which are by now independently confirmed) the antineutrino flux through their detector was $\sim 5 \times 10^{17}$ antineutrinos per second per meter squared (obtained from the Los Alamos reactor), yet only a handful of type (2.22) reactions were registered per hour.

Thus, in 1956 – a quarter of a century after his original proposal – Pauli's insistence on preserving the conservation laws triumphed! As we will see later, conservation laws are directly related to symmetries – which is the content of Amalie Emmy Noether's theorem. Reliance on symmetries and conservation laws was thereby irrevocably infused into the understanding of Nature – which is highly ironic, given Pauli's denigrating attitude towards group theory [☞ p. 150], which however governs the structure of symmetries. We will see that the success of relativistic physics also may be understood from the point of view of symmetries, although this was definitely not evident at the time.

Applying crossing symmetry to the processes (2.20) and (2.22), we know that the process

$$\nu_e + n^0 \longrightarrow p^+ + e^- \tag{2.23}$$

must also occur. To check if the neutrino is its own antiparticle, Raymond Davis, Jr. and Donald S. Harmer then looked for signals of the analogous reaction

$$\overline{\nu}_e + n^0 \longrightarrow p^+ + e^-, \qquad \text{(this does not occur!)} \tag{2.24}$$

and found no signal although the set-up and analysis was analogous to that of the earlier Reines–Cowan experiment. From the fact that the process (2.22) does occur while (2.24) does not, it follows that a neutrino is distinguishable from an antineutrino, and that we may associate with them opposite values of a conserved quantity.

In fact, in 1953, Konopinski and Mahmoud [319] proposed such a conserved quantity. With a small adaptation of their original proposal, we may call this conserved quantity the *lepton number*, so that

$$L = +1: \quad e^-,\ \nu_e; \quad \mu^-,\ \nu_\mu; \quad \tau^-,\ \nu_\tau; \tag{2.25a}$$
$$L = -1: \quad e^+,\ \overline{\nu}_e; \quad \mu^+,\ \overline{\nu}_\mu; \quad \tau^+,\ \overline{\nu}_\tau; \tag{2.25b}$$
$$L = 0: \quad \text{all other particles.} \tag{2.25c}$$

Of course, in 1953 one knew nothing of the existence of the $\tau^\pm$-leptons and tau-neutrinos, and even the existence of muon-neutrinos (as distinct from electron-neutrinos) was not clear. Nevertheless, using the values (2.25), the reactions (2.20)–(2.23) are permitted, whereas (2.24) is forbidden by the lepton number conservation law.

Conservation of the lepton number – as defined by the values (2.25) – is a law that may be used much as electric charge conservation, although the values (2.25) are ascribed to particles so as to explain the occurrence of processes like (2.20)–(2.23) and the absence of processes like (2.24). Just as electric charge conservation is related to a certain continuous $U(1)$ symmetry [☞ Chapter 5.1–5.3], so lepton number conservation has its "own" symmetry, which will be important when discussing the Lagrangian for the theory of lepton interactions.

Finally, none of the so far mentioned conservation laws prevents the potential decay

$$\mu^- \xrightarrow{?} e^- + 2\gamma, \tag{2.26}$$

but this process was never observed. Experience with quantum physics of atomic transitions and reactions and also the increasing number of processes with elementary particles indicates the rule:

> ***Conclusion 2.3*** *(Murray Gell-Mann's "totalitarian principle," cited from Ref. [567])*
> *Everything not forbidden is compulsory.*

Thus, the absence of the decay (2.26) requires an explanation[17] in the form of a proposal that leptons $e^{\pm}$, ν_e and $\overline{\nu}_e$ have a separately conserved number, as do $\mu^{\pm}$, ν_μ and $\overline{\nu}_\mu$, and also the later discovered $\tau^{\pm}$, ν_τ and $\overline{\nu}_\tau$.

The absence of the process (2.26) is thus seen as the manifestation of the ban imposed by the separate muon- and electron-number conservation laws. These laws in turn do permit the decays (2.21a)–(2.21b): the net sum of each of the three separate lepton numbers on the "before" side of either of those processes equals the sum of those same lepton numbers on the "after" side. The same holds for all other listed processes, including Fermi's formula (2.20) for β-decay. This illustrates the basic principle that laws of Nature must have no exception and must hold universally.[18]

This analysis, following the proposal by Konopinski and Mahmoud [319], formally introduced a separate conservation law for the electron-, muon- and tau-lepton numbers, in addition to the conservation laws of electric charge, 4-momentum (i.e., energy and 3-dimensional linear momentum), and of the angular momentum.

— ❦ —

Finally, it is worth noting that the name *lepton* stems from the Greek adjective λεπτός (light), because of the relatively small electron mass, as compared to that of the proton and the neutron. For the latter two, the collective name is *baryon*, from the Greek adjective βαρύς (heavy). Particles with masses between $m_e \approx 0.511$ MeV/c^2 and $m_p \approx 938$ MeV/c^2 were thus named *mesons*, from the Greek word μέσος (middle). However, the discovery of the muon and the verification that it is identical to the electron – except for being ~206 times heavier – suggested that the original and naive nomenclature had to change. By the mid-twentieth century, these names were re-purposed according to the interaction type, as shown in Table 2.2. The fact that the lepton and the baryon numbers are conserved in all processes, whereas the number of mesons is not, is a feature of Nature that slowly became ever clearer, through the analysis of an ever larger number of processes. The name *hadron* stems from the Greek word ἁδρός (thick, bulky).

Table 2.2 Defining collections of elementary particles according to their interactions

	Group	**Nuclear interactions**	**Spin**	**Number**
	Leptons	Only weak	Half-integral	Conserved
Hadrons	**Mesons**	Both strong and weak	Integral	Not conserved
	Baryons	Both strong and weak	Half-integral	Conserved[a]

[a] Baryon number (albeit not under that name) conservation proposed in 1938 by Ernst Stückelberg, to explain the absence of the $p^+ \to e^+ + \pi^0$ proton decay.

2.3.11 Strange particles

By mid-1947, there was perfect experimental proof of the existence of the electron, the proton and the neutron, of which practically all substance around us is composed. Yukawa's π-meson

---

[17] Candidates for a law of Nature are not proven but disproven by exceptions.

[18] One often says that for small speeds non-relativistic physics holds and that for speeds that are near the speed of light in vacuum relativistic physics holds. Literally taken, this is false: what is true is that relativistic physics holds always, but that for small enough speeds the non-relativistic approximations *suffice in practice*. That is, the difference between particular concrete results of relativistic computations and their non-relativistic approximations cannot be experimentally detected.

was also experimentally detected, so that there existed a real chance for a theoretical description of the strong nuclear interactions to be developed so as to adequately reproduce the experimental facts about atomic nuclei. Fermi's theory of β-decay adequately described all known effects of weak nuclear interaction. The antiparticle of the electron that Dirac predicted was also detected experimentally and there was no doubt that, upon appropriate development of experimental devices, all other antiparticles would be experimentally produced. The existence of the neutrino had experimentally still not been verified, but at least ever more theorists agreed that it did have to exist.

Thus, only the existence of the muon presented a capricious puzzle of Nature: this about 206 times heavier copy of the electron was completely unexpected and unexplained.

— ❦ —

In December of 1947, George D. Rochester and Clifford C. Butler opened Pandora's box: They published the results of their analysis of photographs of cosmic rays in the (Wilson) cloud chamber, from which it followed that there existed a neutral particle with a mass of about half the neutron mass, and which decays

$$K^0 \longrightarrow \pi^+ + \pi^- . \tag{2.27}$$

In 1949, Cecil Powell published the experimental discovery of a new charged particle that decays

$$K^+ \longrightarrow \pi^+ + \pi^+ + \pi^- . \tag{2.28}$$

In 1950, Carl D. Anderson's group at CalTech discovered a new neutral particle that decays

$$\Lambda^0 \longrightarrow p^+ + \pi^- . \tag{2.29}$$

Only by 1956 did it transpire that K^0 and K^+ are closely related, just as if they were heavier copies of π^0 and π^+, but even in these early years it was clear that these new particles were rather unusual: $K^0, K^\pm, \Lambda^0$ – and soon several more were discovered – were all produced in very fast ($\sim 10^{-23}$ s) reactions, but their half-life (and lifetime) was relatively long: $\sim 10^{-10}$–10^{-8} s.

It soon turned out that these new particles were created in pairs, so Abraham Pais proposed the concept of "associated production." In 1953, Kazuhiko Nishijima and Tadao Nakano transformed this proposal into a concept of the "eta charge," and in 1965, Murray Gell-Mann independently introduced the "strangeness" charge. Under that name, the idea was finally adopted: When "strange" particles are created, they are created in "strange–antistrange" pairs, which indicates a strangeness conservation law. Thereafter, the decay of a "strange" ($S_0 \neq 0$) particle into a collection of particles the total strangeness of which is $\sum_i S_i \neq S_0$ would be forbidden by this strangeness conservation law, and so could happen only via an interaction that violates this law explicitly. This then establishes that strange particles are created by one (strong) interaction, and decay by another (weak). In addition, this proposal also contained the so-called GNN formula:

$$Q = I_3 + \tfrac{1}{2}(B + S), \tag{2.30}$$

after the initials of Gell-Mann, Nishijima and Nakano. Here, Q denotes the electric charge, I_3 is the isospin [☞ discussion around relations (2.7)–(2.12)], B the baryon number and S denotes the *strangeness*. The fact that these quantum numbers may be consistently assigned to a growing number of particles (both mesons and hyperons[19]) while satisfying the relation (2.30) indicates a regularity that needed an explanation.

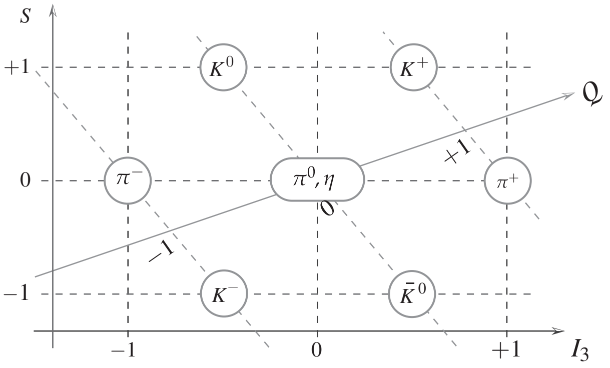
---

[19] The word "hyperon" was initially used for particles heavier than the neutron; nowadays, it is used for strange baryons with neither charm, nor beauty, nor truth [☞ table in display (2.44a)].

2.3.12 The eightfold way

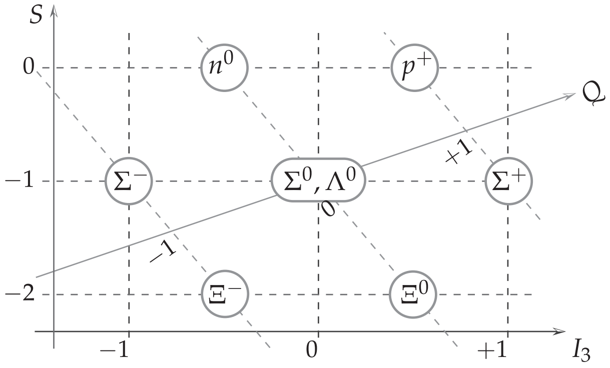
In the early 1930s, the list of elementary particles was short and really simple: Substance consists of electrons, protons and neutrons. These particles interact via electromagnetic forces, mediated by photons, weak nuclear forces formulated by Fermi as a so-called contact interaction (with no mediator), and strong nuclear forces for which Yukawa's theory with pions as mediators seemed a good candidate. The fourth fundamental interaction, gravitation, had by then been described by Einstein's general theory of relativity.

However, by 1960, this list of elementary particles was joined by so many new particles that a systematization became necessary, somewhat akin to Mendeleev's periodic table of elements. Murray Gell-Mann noticed that a 2-dimensional plot of the first eight pseudo-scalar[20] mesons:



$$\begin{aligned} m_{\pi^0} &= 135.0\,\mathrm{MeV}/c^2 \\ m_{\pi^\pm} &= 139.6\,\mathrm{MeV}/c^2 \\ m_{\eta} &= 547.9\,\mathrm{MeV}/c^2 \\ m_{K^\pm} &= 493.7\,\mathrm{MeV}/c^2 \\ m_{K^0} &= 497.6\,\mathrm{MeV}/c^2 \\ \text{spin} &= 0, \end{aligned} \tag{2.31}$$

looks very similar to the analogous plot of baryons:



$$\begin{aligned} m_{p^+} &= 938.3\,\mathrm{MeV}/c^2 \\ m_{n^0} &= 939.6\,\mathrm{MeV}/c^2 \\ m_{\Lambda} &= 1.116\,\mathrm{GeV}/c^2 \\ m_{\Sigma^+} &= 1.189\,\mathrm{GeV}/c^2 \\ m_{\Sigma^0} &= 1.193\,\mathrm{GeV}/c^2 \\ m_{\Sigma^-} &= 1.197\,\mathrm{GeV}/c^2 \\ m_{\Xi^0} &= 1.315\,\mathrm{GeV}/c^2 \\ m_{\Xi^-} &= 1.322\,\mathrm{GeV}/c^2 \\ \text{spin} &= \tfrac{1}{2}. \end{aligned} \tag{2.32}$$

Besides, Gell-Mann had in 1961, and Susumu Okubo independently in 1962, noticed that the masses of the particles in these diagrams satisfy (up to a small percentage error) the relations

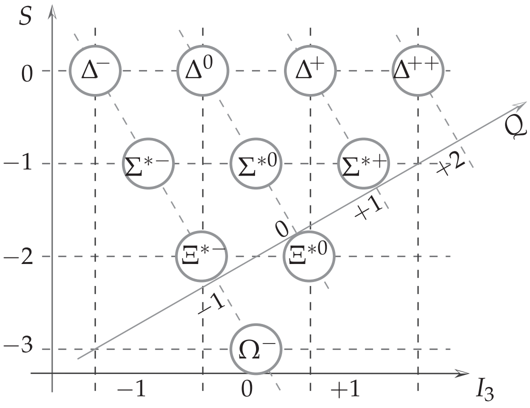
$$2\left[\left(\frac{m_{K^-}+m_{\bar{K}^0}}{2}\right)^2+\left(\frac{m_{K^0}+m_{K^+}}{2}\right)^2\right] \approx 3m_\eta^2+\left(\frac{m_{\pi^-}+m_{\pi^0}+m_{\pi^+}}{3}\right)^2, \tag{2.33}$$

$$2\left[\frac{m_p+m_n}{2}+\frac{m_{\Xi^-}+m_{\Xi^0}}{2}\right] \approx 3m_\Lambda+\frac{m_{\Sigma^-}+m_{\Sigma^0}+m_{\Sigma^+}}{3}. \tag{2.34}$$

---

[20] A scalar function has spin 0, and does not change under rotations; the prefix *pseudo* then indicates that unlike real scalars that are invariant also with respect to the $\vec{r} \to -\vec{r}$ reflection, pseudo-scalars change their sign.

The next collection of baryons forms a somewhat different figure:



$$\begin{aligned} m_\Delta &= 1.231\text{–}1.233\,\mathrm{GeV}/c^2 \\ m_{\Sigma^*} &= 1.383\text{–}1.387\,\mathrm{GeV}/c^2 \\ m_{\Xi^*} &= 1.532\text{–}1.535\,\mathrm{GeV}/c^2 \\ m_\Omega &= 1.672\,\mathrm{GeV}/c^2 \\ \text{spin} &= \tfrac{3}{2}, \end{aligned} \tag{2.35}$$

and where Ω^- has not yet been discovered. Following the approximate relation

$$\frac{9m_\Lambda + m_{\Sigma^-} + m_{\Sigma^0} + m_{\Sigma^+}}{12} - \frac{m_p + m_n}{2} \approx \frac{m_{\Xi^-} + m_{\Xi^0}}{2} - \frac{9m_\Lambda + m_{\Sigma^-} + m_{\Sigma^0} + m_{\Sigma^+}}{12}, \tag{2.36}$$

between the masses of the baryons in the octet (2.32), i.e., that their average masses grow uniformly with the increasing of the absolute value of strangeness,[21] Gell-Mann postulated the relation

$$m_\Delta - m_{\Sigma^*} = m_{\Sigma^*} - m_{\Xi^*} = m_{\Xi^*} - m_\Omega, \tag{2.37}$$

where m_Δ, m_{Σ^*} and m_{Ξ^*} are the average masses of the particles in the upper three rows in the plot (2.35). This predicted the existence of the Ω^- particle with a mass of about $1.70\,\mathrm{GeV}/c^2$, up to a small percentage. In 1964, this particle was experimentally detected, with a mass that differs only 0.6 % from the value predicted by Gell-Mann's relation (2.37). The Reader may have asked why, e.g., four Δ-particles are not in the same collection with the proton, the neutron, and the Λ- and Σ-particles, since the masses of these baryons are closer to the mass of the Δ-particle than to the mass of the Ξ-particle. As it was done in those early 1960s, the answer is pragmatic: because the values of the spin indicated in the plots (2.32) and (2.35), and the formulae (2.34), (2.36) and (2.37) gather particles into collections as shown and not otherwise; [☞ Section 4.4].

This proved the classifying scheme that Gell-Mann called the "eightfold way" (alluding to the "noble eightfold way" of Buddhism). Even though in the 1960s this was not entirely clear, these results crucially use $SU(3)$ symmetry – because of which the plots (2.31), (2.32) and (2.35) are nowadays usually drawn with the S-axis at an angle of 120° with respect to the Q-axis, as is customary for the root system of the $SU(3)$ group [581, 105, 256, 447].

It should be noted that the meson collections, such as (2.31), contain both particles and antiparticles, in diametrically opposite positions: π^- is anti-π^+, K^- is anti-K^+, etc. On the other hand, the baryon groups (2.32) and (2.35) contain only particles, and their antiparticles form identical collections but with charges that all have opposite signs.

2.3.13 Quarks

The true meaning of Gell-Mann's "eightfold way," however, is the systematic use of group theory in particle classification, as was independently proposed by Yuval Ne'eman about the same time. Whereas Mendeleev's periodic table (1869) waited many years for a final explanation (1925) by

---

[21] The sign and an additive constant in the definition of the quantum number of strangeness are for all present purposes completely arbitrary.

way of quantum mechanics and Pauli's exclusion principle, the "eightfold way" was explained as early as 1964: Murray Gell-Mann and George Zweig independently proposed the quark model,[22] whereby mesons are quark–antiquark bound states, and baryons three-quark bound states. This explained all the hadrons in a fast growing list as composite systems, consisting of only three quarks:

Name	$q : Q$	I_3	B	S
Up	$u : +\frac{2}{3}$	$+\frac{1}{2}$	$\frac{1}{3}$	0
Down	$d : -\frac{1}{3}$	$-\frac{1}{2}$	$\frac{1}{3}$	0
Strange	$s : -\frac{1}{3}$	0	$\frac{1}{3}$	-1

(2.38)

the various charges of which satisfy the GNN formula (2.30). These three quarks span the **3** representation of the $SU(3)$ group, and all the hadrons then must form as groupings according to [☞ Appendix A]

$$\text{mesons} = (q\,\bar{q}) \;:\quad \mathbf{3}\otimes\mathbf{3}^* = \mathbf{1}\oplus\mathbf{8}, \tag{2.39}$$

$$\text{baryons} = (q\,q\,q) \;:\quad \mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3} = \mathbf{1}\oplus\mathbf{8}\oplus\mathbf{8}\oplus\mathbf{10}, \tag{2.40}$$

which predicts that mesons must form *singlets* (separate states) or *octets* (collections of eight), whereas baryons must form *singlets* or *octets* or *decuplets*; see Example A.6. This is consistent with experiments: for each collection of particles there exist formulae of the forms (2.33)–(2.34) and (2.36)–(2.37). The combinatorial details of the quark model (and why there are in fact no singlet baryons) are the subject matter of Chapter 4.

Together with the success in classifying hadrons, the quark model also had two important problems:

1. no experiment could produce a free quark;
2. in some baryons it seemed that the existence of the three-quark state violated Pauli's exclusion principle.

The impossibility of extracting *free* quarks does not imply that they were impossible to verify experimentally: One merely needs an adequate probe to "see" the composite structure of a hadron without freeing its constituent building blocks, just as Rutherford bombarded gold atoms with α-particles and "saw" the atomic nuclei without (even partially) ionizing the atoms. In the late 1960s such *deep inelastic* collisions were performed at SLAC (*Stanford Linear Accelerator Center*), bombarding protons with electrons, and later in the 1970s at CERN,[23] bombarding protons with neutrinos, and then also with protons. Much as in Rutherford's experiment, the probes pass right through the protons – in growing number at growing energies and with little deflection, while a small number deflect at a large angle (even to 180°). However, the details of the scattering indicate that the proton is very well represented as a composite system consisting of *three* particles, in agreement with the quark model. Also, while we *can* ionize atoms and so fully liberate their nuclei, quarks cannot be extracted free from the hadrons. Repeated failures to do so created an undercurrent of mistrust in the quark model, whereby most experimentalists rather used Feynman's term, *partons*, for the scattering centers within the hadrons.

---

[22] Interestingly, Gell-Mann nevertheless advocated that the quarks are not necessarily "real," concrete particles in the usual sense with well-localized and detectable position [☞ Digression 4.2 on p. 151]. In this respect, he drastically differed from Richard Feynman (also at CalTech). Feynman advocated for real particle constituents, but called them "partons" to avoid Gell-Mann's quarks. A decade later, Gell-Mann's quarks (together with gluons) proved to be Feynman's partons. On the other hand, Zweig used the term "ace" (presumably alluding to "ace up the sleeve") for quarks, but that never caught on [592, 119].

[23] In 1954, CERN was established as *Conseil Europeen pour la Recherche Nucleaire*, the provisional council that established the laboratory *Organisation Europeenne pour la Recherche Nucleaire*, keeping however the acronym.

On the other hand, in 1964, Oscar W. Greenberg noticed that the quark model seems to violate Pauli's exclusion principle: certain spin-$\frac{3}{2}$ states such as $\Delta^- = (d\,d\,d)$, $\Delta^{++} = (u\,u\,u)$ and the celebrated $\Omega^- = (s\,s\,s)$ should be forbidden owing to Pauli's exclusion principle. In all three of these cases, experiments indicate that the three *otherwise identical* quarks are in the S-state, with parallel spins, and so in the very same quantum state – contradicting Pauli's exclusion principle. Greenberg thus proposed [229] that quarks satisfy *para-fermionic* (anti-)commutation rules: Formally, while bosonic creation and annihilation operators commute and the analogous fermionic operators anticommute,

$$\big[\,b_i\,,\,b_j^{\,\dagger}\,\big] = \delta_{ij}, \qquad \big[\,b_i\,,\,b_j\,\big] = 0 = [b_i^{\,\dagger}, b_j^{\,\dagger}], \qquad \text{bosons,} \tag{2.41a}$$

$$\big\{\,f_i\,,\,f_j^{\,\dagger}\,\big\} = \delta_{ij}, \qquad \big\{\,f_i\,,\,f_j\,\big\} = 0 = \{f_i^{\,\dagger}, f_j^{\,\dagger}\}, \qquad \text{fermions,} \tag{2.41b}$$

para-fermion creation and annihilation operators satisfy the hybrid relations:

$$\left.\begin{aligned}
&\big\{\,\tilde f_{i,\alpha}\,,\,\tilde f^{\dagger}_{j,\alpha}\,\big\} = \delta_{ij}, \qquad \big\{\,\tilde f_{i,\alpha}\,,\,\tilde f_{j,\alpha}\,\big\} = 0 = \big\{\,\tilde f^{\dagger}_{i,\alpha}\,,\,\tilde f^{\dagger}_{j,\alpha}\,\big\}, && (2.41\text{c})\\
&\big[\,\tilde f_{i,\alpha}\,,\,\tilde f^{\dagger}_{j,\beta}\,\big] = \delta_{ij}, \qquad \big[\,\tilde f_{i,\alpha}\,,\,\tilde f_{j,\beta}\,\big] = 0 = \big[\,\tilde f^{\dagger}_{i,\alpha}\,,\,\tilde f^{\dagger}_{j,\alpha}\,\big], \quad \alpha \neq \beta, && (2.41\text{d})
\end{aligned}\right\}\ \text{para-fermions,}$$

where $\alpha, \beta = 1, 2, 3$. In January 1965, Boris V. Struminsky proposed in a paper presented at Dubna (Moscow region, Russia) an "additional" quantum number to resolve this problem, and continued working on this with his mentor Nikolay Bogolyubov, and collaborator Albert Tavchelidze. In May 1965, Tavchelidze presented this idea at ICTP, in Trieste (Italy), without his collaborators' knowlege. Six months later, Moo-Young Han and Yoichiro Nambu independently proposed a model where a new degree of freedom, α, β in (2.41c)–(2.41d) is a new kind of "charge" and has its own interaction with eight new "photons," $g_\alpha{}^\beta$ where $g_\alpha{}^\alpha = 0$. This is what is called *color* today, and which differentiates the quarks in the hadrons $\Delta^-, \Delta^{++}, \Omega^-$ so Pauli's exclusion principle would not forbid their existence. In their model, quarks had integer, but color-dependent electric charges [☞ Digression 5.14 on p. 214]. The final version of the formalism of *color* as a charge for strong interaction and which is independent from the electric charge was completed by William Bardeen, Harald Fritzsch and Murray Gell-Mann, in 1974.

On the third hand, in 1964, Sheldon L. Glashow and James D. Bjorken had proposed the existence of a fourth quark, dubbed *charm*, and with it an extension of the classification $SU(3)$ group into $SU(4)$. In 1970, Glashow, John Iliopoulos and Luciano Maiani provided a theoretical proof that the fourth quark must exist based on the absence of weak decays that would change strangeness:

$$\frac{K^+ \to \pi^+ + \bar{\nu} + \nu}{K^+ \to \pi^0 + \mu^+ + \mu^-} = \frac{(u\bar{s}) \xrightarrow{Z^0} (u\bar{d}) + \bar{\nu} + \nu}{(u\bar{s}) \xrightarrow{W^+} (u\bar{u}) + \mu^+ + \mu^-} < 10^{-5}. \tag{2.42}$$

Besides, the appearance of the so-called Adler–Bell–Jackiw (ABJ) $U(1)$ *anomaly* [☞ Section 7.2.3, and the lexicon entries about anomaly and canonical quantization] indicates an inconsistency in the gauge theory of weak interaction: the quark model with only three quarks exhibits a symmetry that is broken by quantum effects, but this breaking cancels through contributions from the fourth quark. That is the first application of a detailed quantum analysis of symmetries; this was later developed into a very powerful theoretical method and gave rise to certain exact (non-perturbative) results in field theory. The c-quark was experimentally discovered four years later. However, a more detailed analysis requires details of both gauge theories and of left–right asymmetric weak interactions, and we will return to this topic in Sections 7.2.2 and 7.2.3.

Finally, in 1973, Makoto Kobayashi and Toshihide Maskawa showed that the so-called *indirect* CP-violation – which James W. Cronin and Val L. Fitch observed back in 1964 (and for which they received the Nobel Prize as late as 1980) in the unevenness of the $K^0 \leftrightarrow \bar{K}^0$ transmutation – can happen only if there exist at least six quarks. The direct *CP*-violation was observed in the 1990s, in agreement with the Kobayashi–Maskawa proposal. By that time the fifth, b-quark had been experimentally produced, while the sixth, t-quark was produced five years later, in 1995. The species of quarks

$$u\ (\text{up}),\quad d\ (\text{down}),\quad s\ (\text{strange}),\quad c\ (\text{charm}),\quad b\ (\text{bottom}),\quad t\ (\text{top}), \tag{2.43}$$

are dubbed *flavors*, so that the symmetry group that arises from approximate identification of these quarks: $SU(3)$, $SU(4)$, $SU(5)$ and $SU(6)$ is typically labeled by the subscript "f." These $SU(n)_f$ approximate "flavor" symmetries are less and less practical for larger and larger n. The quark masses are more and more different, and the symmetries are less and less precise:

Name	q	**Mass*** (MeV/c^2)	Q	I_3	B	S	C	B'	T	Y
Up	u :	1.5–3.3	$+\frac{2}{3}$	$+\frac{1}{2}$	$\frac{1}{3}$	0	0	0	0	$+\frac{1}{3}$
Down	d :	3.5–6.0	$-\frac{1}{3}$	$-\frac{1}{2}$	$\frac{1}{3}$	0	0	0	0	$+\frac{1}{3}$
Strange	s :	$105\{^{+25}_{-35}$	$-\frac{1}{3}$	0	$\frac{1}{3}$	-1	0	0	0	$-\frac{2}{3}$
Charm	c :	$1{,}270\{^{+70}_{-110}$	$+\frac{2}{3}$	0	$\frac{1}{3}$	0	$+1$	0	0	$+\frac{4}{3}$
Bottom	b :	$4{,}200\{^{+170}_{-70}$	$-\frac{1}{3}$	0	$\frac{1}{3}$	0	0	-1	0	$-\frac{2}{3}$
Top	t :	$171{,}300\{^{+1{,}100}_{-1{,}200}$	$+\frac{2}{3}$	0	$\frac{1}{3}$	0	0	0	$+1$	$+\frac{4}{3}$

(2.44a)

* Inertial mass without the binding energy, which depends on the hadron

$$Q = I_3 + \tfrac{1}{2}\underbrace{(\mathit{Baryon} + \mathit{Strange} + \mathit{Charm} + \mathit{B'eauty} + \mathit{Truth})}_{=Y,\ \text{so-called (strong) hypercharge [☞ Section 7.2.1]}} \tag{2.44b}$$

During 1964–74, feelings about the quark model were rather mixed. While experimentalists rightfully decried the fact that quarks seemed to be impossible to extract free – for which there was no theoretical explanation – even the quantum number of color, introduced to reconcile the quark model with Pauli's exclusion principle, seemed more of a mnemonic crutch than a real physical property. Just as the quarks seemed impossible to extract, so was the color impossible to detect directly: The three quarks in a baryon each have a different color, red–blue–yellow, so that baryons are "colorless." The quark and the antiquark in a meson have opposite colors (red–green, blue–orange, or yellow–purple), so that mesons are also "colorless." As a classifying system, this rule perfectly predicted hadronic states [☞ Chapter 4]. However, the skeptical physicists could not escape the impression that the color formalism was "invented" so as to "explain" the otherwise unexplained fact: the formalism gave no *reason* for hadrons to be "colorless" [☞ Section 6.1].

Finally, note that the quark model revised the elementary particle image: all substance consists of quarks (u, d, s, c, b, t) and leptons $(e^-, \nu_e, \mu^-, \nu_\mu, \tau^-, \nu_\tau)$, which returns simplicity into the list of elementary constituents of the World: by 1974, the number of hadrons had approached a hundred and no one could possibly consider them elementary. By contrast, the list of (then known) quarks and leptons was short and even fairly "symmetric."

In spite of this *theoretical* and *aesthetic* attractiveness, the event that was crucial in winning the confidence of most physicists in the quark model was the *experimental* detection of the so-called J/ψ particle, towards the end of 1974. During the summer of 1974, C. C. Samuel Ting's research group in Brookhaven discovered a 3.096 9 GeV/c^2-mass particle (a little over three times the proton mass), and with a half-life measured to be around 10^{-20} s – which is some 1,000 times longer than the typical hadron half-life! Ting insisted on careful checking of this astounding result, so that the discovery remained a secret till November of that year, when Burton Richter's research group at SLAC discovered the same particle, and the two research groups published their results back-to-back.

Over the next few months, it became clear that the J/ψ particle was a $(c\,\bar{c})$ bound state, i.e., a meson,[24] which, being akin to positronium, is called "charmonium." The stability of this system follows from the combination of the so-called OZI rule (Susumu Okubo, George Zweig and Jugoro Iizuka) and the fact (discovered soon) that the pair of lightest mesons with a single c- or a single $\bar{c}$-quark (the so-called D, i.e., $\bar{D}$ meson, respectively) is heavier than the J/ψ particle. Thus, a J/ψ does not have enough mass to decay into a D–$\bar{D}$ pair, and the decays that require the c–$\bar{c}$ annihilation are slowed down by the OZI rule.[25] When the J/ψ particle was discovered, it became obvious that its existence fits so perfectly in the quark model that all doubt in the model vanished: As should be clear from the foregoing, and also Table 2.4 on p. 69, the fourth quark was already predicted, both from aesthetic (Bjorken and Glashow) and also technical (what was only later understood as most stringent and rigorous) reasons of anomaly cancellation (Glashow, Iliopoulos and Maiani).

In 1975, a new lepton, τ^-, was discovered, supporting the prediction (Kobayashi and Maskawa, two years earlier) of another lepton and quark pair, so as to explain *CP*-violation. Just as with the J/ψ particle, a $(b\,\bar{b})$ bound state was discovered in 1977 and dubbed Υ. The sixth, t-quark was finally detected in 1995, and the τ-neutrino, ν_τ, in 2000. Thus, by the dawn of the third millennium, the list of elementary particles (2.44a) was experimentally confirmed.

Digression 2.5 It should be noted that the charming story with charmonium, and even with *bottomonium* (i.e., with the $(b\,\bar{b})$-state Υ) will probably not be repeated with *toponium*: The t-quark itself has a 174.2 GeV/c^2 mass, so that according to Ref. [215] and contemporary data [293] the $(t\,\bar{t})$-state would have to have a mass of about 344.4 GeV/c^2, and a standard deviation of $\sigma \approx$ 193.6 MeV/c^2. Thus, even when the experiments reach energies over about 344.4 GeV, the large absolute value of the standard deviation implies that the toponium will decay much faster than the J/ψ and the Υ [☞ Section 2.4], and it may well turn out that it will behave practically as a virtual particle, which by definition cannot be detected directly.

2.3.14 Nuclear force intermediaries

The introduction of *color*, the new quantum number assigned to quarks but not to leptons, not only explained the existence of spin-$\frac{3}{2}$ baryons such as $\Delta^-, \Delta^{++}, \Omega^-$ in terms of three-quark S-state bound states, but also brought about the ultimate explanation of the strong nuclear interaction. The exchange of color between quarks is mediated by gluons, which thus mediate the strong interaction; the details of this mechanism will be examined in Chapters 5–6. Suffice it to say, the theoretical basis for the so-called Yang–Mills theory of non-abelian (non-commutative) gauge symmetry had been introduced back in 1954–5: Chen-Ning Yang and Robert L. Mills and, independently, Ronald Shaw in his PhD dissertation under Abdus Salam, showed how the electromagnetic

---

[24] By 1974, it became evident that the old motivation for the nomenclature, whereby *mesons* had masses between those of the electron and the proton, was no longer practical. Instead, the quark model convention was adopted, wherein every three-quark bound state was called a *baryon*, and every quark–antiquark bound state was called a *meson*. Hypothetical states of other composition, such as $(q\,q)$, $(q\,q\,\bar{q})$, $(q\,q\,q\,q)$, etc., were referred to as exotic particles.

[25] A decade later, the OZI rule was derived from quantum chromodynamics, but in 1974 it was still only a phenomenological *rule*.

interaction with $U(1)$ symmetry group [☞ Sections 5.1–5.3] can be generalized into a gauge interaction with a non-abelian (non-commutative) symmetry [☞ Section 6.1, as well as Appendix A]. However, the crucial (and in fact complementary) qualities of the quark model,

1. that quarks cannot be extracted free from hadrons (so-called confinement), and
2. that the closer quarks are to each other, the weaker they interact (so-called asymptotic freedom, experimentally observed as early as 1967)

are not obvious consequences of the model. This latter quality was discovered by David Gross with his student Frank Wilczek, and independently by David Politzer in 1973,[26] almost two decades after the discovery of non-abelian gauge theory itself. Although the theoretical proof of the first quality (confinement) is still not rigorously complete☝, the proof of asymptotic freedom as its complementary quality caused an overnight universal acceptance of quantum chromodynamics (QCD) as *the* theory of strong nuclear interactions. In addition, numerical computations in so-called “lattice QCD” (where the infinite and continuous spacetime is approximated by a finite-size lattice with a nonzero lattice spacing) and “Monte Carlo simulations” soon showed that QCD correctly reproduces many of the ratios of hadron masses as well as many other parameters of so-called hadron spectroscopy, so that the crucial unsolved problem☝ is “only” to compute the absolute value of a characteristic mass unit for hadrons. It is worth noting that the theoretical discovery of asymptotic freedom and the experimental discovery of the J/ψ particle practically coincided. Without a doubt, that tandem advancement was decisive in the sudden turn of tide in accepting the quark model together with quantum chromodynamics, and this combination of events is sometimes referred to as the “November Revolution” of 1974.

* * *

Meanwhile, for weak nuclear interactions, Enrico Fermi formulated the so-called 4-fermion contact interaction in 1931–4, which within the quark model stated is as

$$d \to u + e^- + \overline{\nu}_e, \tag{2.45a}$$

together with all possible related processes obtained via crossing symmetry and the principle of detailed balance (2.14), such as

$$d + \overline{u} \to e^- + \overline{\nu}_e, \qquad d + e^+ \to u + \overline{\nu}_e, \qquad u + e^- \to d + \nu_e, \qquad \text{etc.} \tag{2.45b}$$

Since the range of the weak interactions is small ($\sim 10^{-15}$ m), Fermi’s approximation, where the weak interaction happens in a point, is quite satisfactory up to energies of about $\hbar c/(10^{-15}\,\text{m}) \sim 200\,\text{MeV}$. However, within two decades after Fermi’s theory, such energies had been surpassed in accelerators and a better theory was needed. The contact interaction (2.45b) is analogous to describing a scattering such as of a positron and a proton, $e^+ + p^+ \to e^+ + p^+$, by neglecting the repulsive interaction field and pretending that the two like-charged particles actually touch during the collision. Akin to electromagnetic interactions, these processes may be described also by introducing *intermediaries*:

---

[26] There is a sad story tied to this discovery [366]: David Gross mentored two students, Frank Wilczek and William E. Caswell. Gross asked Wilczek to compute the so-called β-function for the $SU(n)$ gauge theory to first order in perturbation theory and Caswell to second. When Wilczek discovered that the β-function had the opposite sign from the abelian case, Gross realized the fantastic importance of the result – the theoretical proof that non-abelian gauge theory guarantees asymptotic freedom – and published this immediately with Wilczek. This is one of the most cited papers in the second half of the twentieth century; Gross and Wilczek shared with Politzer, the 2004 Nobel Prize. Caswell’s contribution, which confirmed Wilczek’s and correctly showed that second-order perturbations do not spoil the newly discovered asymptotic freedom was hardly noticed. William E. Caswell died in the Pentagon crash of the American Airlines Flight 77, on September 11, 2001.

$$\begin{aligned} d &\to u + W^- &&\to u + (e^- + \bar{\nu}_e), && (2.46a)\\ d + \bar{u} &\to W^- &&\to e^- + \bar{\nu}_e, && (2.46b)\\ d + e^+ &\to (u + W^-) + e^+ &&\to u + \bar{\nu}_e, && (2.46c)\\ &\searrow d + (W^+ + \bar{\nu}_e) &&\to u + \bar{\nu}_e, && (2.46d)\\ u + e^- &\to (d + W^+) + e^- &&\to d + \nu_e, && (2.46e)\\ &\searrow u + (W^- + \nu_e) &&\to u + e^-, && (2.46f) \end{aligned}$$

and so on. These examples make it clear that we have postulated the intermediaries $W^\pm$ for the weak interaction. Their elementary processes,

$$d + \bar{u} \leftrightarrow W^-, \qquad \bar{d} + u \leftrightarrow W^+, \tag{2.47}$$

$$e^- + \bar{\nu}_e \leftrightarrow W^-, \quad e^+ + \nu_e \leftrightarrow W^+, \tag{2.48}$$

as well as all related processes obtained using the crossing symmetry and the principle of detailed balance (2.14), and by replacing the u, d quarks with the heavier pair c, s (and also t, b), as well as by replacing the e^-, ν_e lepton pair with μ^-, ν_μ (and also τ^-, ν_τ) might seem amply sufficient to describe all known examples of weak interaction. However, that would be false: From the observed decay

$$K^0 = (d\,\bar{s}) \to \pi^+ + \pi^- = (u\bar{d}) + (d\,\bar{u}) \tag{2.49}$$

it follows that the weak elementary processes

$$s + \bar{u} \to W^-, \quad \bar{s} + u \to W^+ \tag{2.50}$$

must also exist, which then also explains the decay

$$\Lambda^0 = (u\,d\,s) \to (u\,d\,W^-\,u) \to (u\,u\,d) + W^- \to (u\,u\,d) + (d\,\bar{u}) = p^+ + \pi^-, \tag{2.51}$$

and so on. Comparing the elementary processes (2.47) and (2.50), it follows that the weak decays “mix” the d- and the s-quark and so violate the conservation of strangeness. That is, the true eigenstates of the Hamiltonian terms responsible for weak interactions[27] are not the $u, d, s, \dots$ quarks, but the combinations

$$|u\rangle, \quad |d_w\rangle := \cos(\theta_c)|d\rangle + \sin(\theta_c)|s\rangle, \quad |s_w\rangle := \cos(\theta_c)|s\rangle - \sin(\theta_c)|d\rangle, \quad \dots \tag{2.52}$$

This was proposed in 1963 – before the discovery of the c-quark – by Nicola Cabibbo and after whom the angle θ_c was named. In 1973, Kobayashi and Maskawa generalized this parametrization by proposing

$$\begin{bmatrix} |d_w\rangle \\ |s_w\rangle \\ |b_w\rangle \end{bmatrix} := \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{bmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{bmatrix}, \tag{2.53a}$$

$$= \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}\,e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}\,e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}\,e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}\,e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}\,e^{i\delta_{13}} & c_{23}c_{13} \end{bmatrix} \begin{bmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{bmatrix}, \tag{2.53b}$$

$$\text{where} \qquad c_{ij} := \cos(\theta_{ij}), \qquad s_{ij} := \sin(\theta_{ij}), \quad i, j = 1, 2, 3 = d, s, b,$$

---

[27] The Student is expected to remember how one computes with Hamiltonians of the form $H = H_0 + H'$ in quantum mechanics, where the eigenstates and eigenvalues for H_0 are known, where H' is treated as a perturbation, and where the eigenstates of H_0 need not be the eigenstates of H', i.e., $[H_0, H'] \neq 0$.

where in the second row the now standard parametrization is given in terms of Euler angles in the (d, s, b)-space. The general form and parametrization of non-Hermitian matrices were known back in 1939 [151], from which it follows that a non-Hermitian $n \times n$ matrix has $\binom{n-1}{2}$ complex phases. It follows that one needs at least three quarks of $-\frac{1}{3}$ electric charge for the Cabibbo–Kobayashi–Maskawa (CKM) matrix to be able to contain one non-removable complex phase, which then can parametrize *CP*-violation, as observed back in 1964. The elements of the matrix $\mathbb{V}$ (2.53) are denoted to indicate their application:

$$\text{Probability}\big(d + W^+ \to u\big) = \big|\langle u|W^+|d\rangle\big|^2 \propto |V_{ud}|^2, \tag{2.54a}$$

$$\text{Probability}\big(s + W^- \to u\big) = \big|\langle u|W^+|s\rangle\big|^2 \propto |V_{us}|^2, \tag{2.54b}$$

and so on. Present-day observed values are $\theta_{12} = \theta_{ds} = (13.04 \pm 0.05)^\circ$, $\theta_{13} = \theta_{db} = (0.201 \pm 0.011)^\circ$, $\theta_{23} = \theta_{sb} = (2.38 \pm 0.06)^\circ$, and $\delta_{13} = \delta_{db} = (1.20 \pm 0.08)^\circ$, giving

$$\begin{bmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{bmatrix} \approx \begin{bmatrix} 0.974 & 0.226 & 0.004 \\ 0.226 & 0.973 & 0.041 \\ 0.009 & 0.041 & 0.999 \end{bmatrix}. \tag{2.55}$$

To estimate the $W^\pm$ particle mass, we need an estimate for the range of weak nuclear forces, and such a direct estimate does not exist. It is known, however, that the weak nuclear interaction does occur within the atomic nucleus in the form of the β-decay, but it is not known how close the particles must be for the scattering to also include weak interactions. For example, in antineutrino–proton scattering, the contribution of the weak interaction would be seen also as the inelastic collision (2.22), which the quark model represents as the consequence of two alternative collisions:

$$\begin{tikzcd}
& (e^+ + W^-) + (u\,u\,d) \arrow[dr] & \\
\overline{\nu}_e + p^+ = \overline{\nu}_e + (u\,u\,d) \arrow[ur] \arrow[dr] & & e^+ + (u\,d\,d) = e^+ + n^0. \\
& \overline{\nu}_e + \big(W^+ + (u\,d\,d)\big) \arrow[ur] &
\end{tikzcd} \tag{2.56}$$

The range of the weak nuclear interaction mediated by $W^\pm$ is thus probably not larger than the diameter of the nucleus where the process takes place, and may well be (much) smaller than the proton and neutron diameter. Taking $R < 10^{-15}$ m for the range produces $m_W > \frac{\hbar}{Rc} \sim 200\,\text{MeV}/c^2$, which is a *lower limit*, and very weak as an estimate: It was known by the late 1940s that no appropriate particle of such a mass exists. As the experiment energies grew, it was expected that the scatterings would begin to show traces of the intermediary bosons $W^\pm$, but such data would be obtained only in January of 1983 (and the Z^0 particle by mid-1983), for which Carlo Rubbia and Simon van der Meer were to receive the 1984 Nobel Prize.

By about 1958, the possibility was noted that there might exist weak neutral processes of the type

$$q + \overline{q'} \leftrightarrow Z^0, \quad \text{and} \quad \ell + \overline{\ell} \leftrightarrow Z^0, \tag{2.57}$$

where q and q' are any two different quarks of the same electric charge – revealing the "mixing" parametrized by the CKM matrix (2.53). Such processes would also produce a correction of the electromagnetic interactions based on the elementary processes

$$q + \overline{q} \leftrightarrow \gamma \quad \text{and} \quad \ell + \overline{\ell} \leftrightarrow \gamma, \tag{2.58}$$

without the CKM mixing and which, of course, do not include the neutrinos as they are electrically neutral. In the processes where q and q' are not the same quark, such as

$$s \to Z^0 + d, \tag{2.59}$$

the quark "flavor" changes, and such hypothetical processes were dubbed "flavor-changing neutral currents" (FCNC). The experimental detection of such processes was crucial for confirming the Glashow–Weinberg–Salam model of electroweak interaction unification based on the $SU(2) \times U(1)$ symmetry group, and the refutation of the competing model based on the $SO(3)$ symmetry group, which was proposed by Sheldon Glashow,[28] developing the idea of his mentor, Julian Schwinger, and with later collaboration by Howard Georgi. Based on this crucial confirmation of their weak interaction model, Glashow, Salam and Weinberg were awarded the 1979 Nobel Prize – five years before the direct detection of $W^{\pm}$ and Z^0 bosons!

The point is that the proof of existence of the FCNC processes sufficed to establish the existence of a Z^0 virtual particle, which then exists only during a time shorter than that estimated by Heisenberg's indeterminacy relations $\triangle t < \frac{\hbar}{2\, m_Z\, c^2}$, and which was therefore not yet directly detected at the time. Such processes can happen (and were detected for the first time in 1973, at CERN) even when the total energy of the collision is not enough to create a "real" Z^0 particle, which then could have been detected directly. It took several years to "improve the statistics" of the results, i.e., to remove the "noise" of the much stronger electromagnetic interaction: In collisions of two particle beams, most processes occur via the much stronger electromagnetic interaction. The probability of identifying a "true" individual process for analyzing the weak interaction in the sea of electromagnetic processes is then very small and requires ingenious technique and methodology in detection as well as an enormous investment in the form of patience.

2.3.15 The Standard Model

By the mid-1980s, the universally accepted list of elementary particles was as given in Table 2.3, and presents the so-called *Standard Model* in its most succinct form. The subsequent chapters will discuss the details of this model (and there are plenty!), but let us note here that the listed 12 spin-$\frac{1}{2}$ particles also have their antiparticles, and that every quark, in addition, has the additional degree of freedom called color, with three distinct values. Thus some Authors [243] count 12 leptons and 36 quarks in Table 2.3. Of course, since these are spin-$\frac{1}{2}$ particles, one should also count the fact that each one of these 48 particles has spin projections $\pm\frac{1}{2}$, which may be regarded as a *doubling* in counting "particles." Similarly, the photon is usually regarded as a single particle, but one must know that every photon has two possible polarizations, which according to this logic should be counted as *two photons*. For nuclear interaction mediators, this number is bigger: For weak interactions there are three intermediary bosons: $W^{\pm}, Z^0$ and each has *three* polarizations,[29] which then gives $3 \times 3 = 9$ particles. Section 6.1 will show that there are 8 gluons and that

Table 2.3 The content of the Standard Model of elementary particle physics; see equation (2.44a)

Substance	(spin-$\frac{1}{2}$ fermions)			
Gen.	**Leptons**		**Quarks**	
1.	ν_e	e^-	u	d
2.	ν_μ	μ^-	c	s
3.	ν_τ	τ^-	t	b

	Interactions		(bosons)
γ $W^{\pm}, Z^0$	electromagnetic weak nuclear	interaction	(spin 1)
gluons	strong nuclear interaction		(spin 1)
$\delta g_{\mu\nu}$	gravitation		(spin 2)

Higgs boson: gives mass to the particles with which it interacts (spin 0)

---

[28] Yes, it is the same Sheldon Lee Glashow, a coauthor of both competing proposals [☞ Footnote *12* on p. 276].

[29] Since the $W^{\pm}, Z^0$ particle mass is not zero, these move with a speed smaller than the speed of light, and have a longitudinal polarization. This is unlike with photons that move at the speed of light so the amplitude of the longitudinal polarization reduces to zero owing to FitzGerald–Lorentz contraction.

they are massless and so have two polarizations each: that adds $8 \times 2 = 16$ particles. Table 2.3 includes gravity, although it is, strictly speaking, not part of the Standard Model; the gravitational field quanta are represented by fluctuations of the metric, and so by a rank-2 tensor. However, those fluctuations propagate at the speed of light and have no mass, and so again have only two polarizations [☞ Chapter 9].

Finally, an integral part of the Standard Model is also the Higgs scalar, which has now been confirmed experimentally [293], the detection of which was one of the original goals of the LHC (Large Hadron Collider) at CERN. Chapter 7 will show that a single real, scalar (spin-0) elementary Higgs particle is predicted, which must be its own antiparticle. While the photon γ, the weak intermediaries $W^{\pm}$ and Z^0, the gluons and even the gravitons are mediating quanta of fundamental interactions, the one real Higgs particle, which has been detected, is a *remnant*: Chapter 7 will show that the Higgs field has *four* real degrees of freedom, three of which are Goldstone modes for spontaneously broken $SU(2)_L$ symmetry. The practical role of the Higgs field is to mediate in giving masses to particles, including the mediating gauge bosons $W^{\pm}, Z^0$ – as if the Higgs field were to slow the particles with which it interacts, reminiscent of the effect of viscosity in materials.

The Goldstone modes in the Higgs field cannot be detected as separate particles, but they can be identified as the additional longitudinal polarizations of the $W^{\pm}$ and Z^0 gauge bosons: In the phase without symmetry breaking there are two complex degrees of freedom of Higgs particles (which may be counted as 2×2 real particles) and three massless gauge bosons of the weak interaction – and so with two polarizations each tallying up to 3×2 real particles; together, that's 10 real particles. In the phase with broken symmetry there is only one real Higgs particle and three massive gauge bosons (three polarizations each), tallying up to 3×3 real particles; together, that's again 10 real particles.

The final sum in this detailed counting is

$$\begin{array}{rcrcr} \text{fermions} & = & (3\times2\times2\times2)+(3\times2\times2\times2)\times3 & = & 96, \\ \text{bosons} & = & 1\times2+3\times3+8\times2+(1\times2)+1 & = & 30, \\ \cline{3-5} & & & = & 126. \end{array} \tag{2.60}$$

In some ways, this *is* the correct counting – and we will see subsequently in what sense one needs to distinguish all these degrees of freedom, as there are physical observables that depend on this level of detail. However, I should like to hope it is clear to the Reader that the complaint "a system of 126 particles does not look elementary" is not fair: It is crucial that these 126 degrees of freedom are systematically presented in the simple Table 2.3. Finally, the particles listed in that table fully explain the by now many hundreds of experimentally detected mesons and baryons, and so all experimentally detected forms of substance (atoms, molecules, etc.), while they themselves show no sign of compositeness or structure. Therefore, the so-called Standard Model with the business card in Table 2.3, and described in more detail in subsequent chapters, fully satisfies the goal of our original search.

Table 2.4 lists a telegraphic review of the most prominent elementary particle physics milestone discoveries.

❦

Motivated mostly by the economy of symmetries in Table 2.3 on p. 67, and a little also by the large number of degrees of freedom (2.60), as early as 1974, there were classification systems that in various ways represent at least some of the 126 particles (2.60) as bound states of even more elementary constituents. Generally speaking, in these proposals quarks (and sometimes also the leptons, and/or the mediating bosons) are bound states of *preons*. Different preon models suppose different dynamics, and then also different combinatorial rules for preons, all with the aim to faithfully reproduce the contents of Table 2.3, and the counting (2.60). However, except for economy

Table 2.4 A timeline of significant discoveries in elementary particle physics

Year	**Particle**	**Discovered**
1895	X-rays	Wilhelm C. Röntgen (X-rays were later identified as photons)
1897	e^-	Joseph J. Thomson
1899	α-particle	Ernest Rutherford
1900	γ-rays	Paul Villard (γ-rays were later identified as photons)
1911	Atomic nucleus	Hans Geiger and Ernest Marsden, under Ernest Rutherford
1919	p^+	Ernest Rutherford
1932	n^0	James Chadwick
1932	e^+	Carl D. Anderson (predicted by Paul A. M. Dirac, 1927)
1937	μ^-	Seth H. Neddermeyer and Carl D. Anderson, Jabez C. Street and Edward C. Stevenson (erroneously identified as pion until 1947)
1947	$\pi^\pm, \pi^0$	Cecil Powell (predicted by Hideki Yukawa, 1935)
1947	K^0	George D. Rochester and Clifford C. Butler
1949	$K^\pm$	Cecil Powell
1947–1953	$\Lambda^0, \Sigma^\pm, \Sigma^0$	Several research groups
1955	$\overline{p}^-, \overline{n}^0$	Owen Chamberlain, Emilio Segrè, Clyde Wiegand and Thomas Ypsilantis
1956	ν (directly)	Frederick Reines and Clyde Cowan (predicted by Wolfgang Pauli, 1931)
1962	$\nu_\mu \neq \nu_e$	Leon M. Lederman, Melvin Schwartz and Jack Steinberger
1969	Partons, and u, d, s quarks	So-called deep inelastic collisions, SLAC (predicted by Murray Gell-Mann and George Zweig, 1963)
1974	J/ψ ($[c\text{-}\bar{c}]$-state)	Burton Richter and C. C. Samuel Ting, proof of the c-quark existence (predicted by James D. Bjorken and Sheldon L. Glashow in 1964, and Glashow, John Iliopoulos and Luciano Maiani in 1970)
1975	τ-lepton	Martin Perl and collaborators, SLAC
1977	Υ (upsilon) ($[b\text{-}\bar{b}]$-state)	Leon Lederman and collaborators, Fermilab (b-quark predicted by Makoto Kobayashi and Toshihide Masakawa in 1973)
1979	Gluon	e^-–e^+ collisions in PETRA experiment at DESY
1983	$W^\pm, Z^0$	Carlo Rubbia, Simon van der Meer, CERN UA-1 collaboration (predicted by Sheldon L. Glashow in 1963, Abdus Salam and Steven Weinberg in 1967)
1995	t-quark	Tevatron, Fermilab (predicted by Makoto Kobayashi and Toshihide Masakawa in 1973)
2000	ν_τ	DONUT collaboration, Fermilab
2012	Higgs	ATLAS and CMS collaborations, LHC at CERN – pending interaction details [25, 109] ☝

in classification and “purely aesthetic” advantages, there exists no *experimental* reason for preon models. In all experiments thus far (at distances $\geqslant \hbar c/E$, where $E \approx 250\,\text{GeV}/c^2$ is the maximal collision energy), quarks, leptons and gauge bosons behave as ideal point-like (elementary) particles. That is, they show no internal structure [☞ Conclusion 1.5 on p. 30]. Of course, the absence of a proof of a structure within quarks and leptons is not a proof of the absence of such a structure.

2.4 Lessons

During the twentieth century the quantumness and relativity of Nature became universally accepted basic ideas of fundamental physics. The third idea that had similarly taken root in our understanding of Nature is the fundamental role of symmetry. On one hand, this links symmetries and conserved quantities: in the form of Amalie Emmy Noether’s theorem in classical physics, rather directly in quantum mechanics, and via so-called Ward–Takahashi identities for gauge theories in quantum field theory. On the other hand, this also links symmetries and interactions, in the form of the gauge principle. Especially in the second half of the twentieth century, symmetries and the algebraic structure of *groups* that those symmetries form so focused the research in fundamental physics that some philosophers of natural sciences [533] acquired the impression that the concept of *law* had been abandoned for symmetry principles. However, the point of view of those who actually do such research is that symmetry groups provide the much needed cohesive (algebraic) structure both to various conservation laws and to the modes of interaction. Symmetries have thus become an integral part of the contemporary understanding of laws of Nature.[30] This difference in understanding the link between symmetries and the laws of Nature reminds us of the comments in Digression 1.1 on p. 9.

The link between symmetry and interaction in the form of gauge theory will be explored in detail in subsequent chapters. Here, we review the conservation laws: We reconsider the logic and rules of their application, provide some cautionary remarks about that application, and list the conservation laws as they are used in elementary particle physics.

2.4.1 *The logic and rules of application*

The growing majority of particles that are studied in elementary particle physics – the hundreds of mesons and baryons – actually are not elementary, but are bound states of quarks and/or antiquarks. All hadrons (except the proton) as well as the $\mu^{\pm}$- and $\tau^{\pm}$-leptons decay, and rather fast [293]: A free n^0 in 15 min, $\mu^{\pm}$ decays in 2.2×10^{-6} s, π^+ in 2.6×10^{-8} s, π^0 in 8.4×10^{-17} s, the J/ψ-particle in 7.05×10^{-21} s, and most of the hadrons decay in a time of about 10^{-23} s! From all those hundreds of particles, only the electron and the proton are stable in the traditional sense – the decay of not one of these has ever been observed.

During 10^{-23} s, light passes only about 3×10^{-15} m in vacuum, which is the order of magnitude of the diameter of the atomic nucleus. It is clear that such short decay lifetimes cannot be measured with a stop-watch. Instead, we use the indeterminacy relation, $\triangle E \, \triangle \tau \geqslant \frac{1}{2}\hbar$, so that $\triangle E = (\triangle m)c^2$ implies

$$\overline{\tau} := \frac{\hbar}{(\triangle m)c^2}, \qquad t_{1/2} := \ln(2)\,\overline{\tau} = \frac{\ln(2)\,\hbar}{(\triangle m)c^2}. \tag{2.61}$$

From the experimentally obtained distribution of the values of the particle mass, one computes the standard deviation of mass and uses it as $\triangle m$ in equation (2.61). The so-computed average duration time of the particle (state), $\overline{\tau}$, is used as the particle lifetime. Then, $t_{1/2}$ is its half-life: $N(t) = N(0)\,e^{-t/\overline{\tau}} = N(0)\,2^{-t/t_{1/2}}$ is the number of a certain type of particles at the time $t > 0$ in a sample where the number at the time $t = 0$ was $N(0)$.

Also, many hadrons (and leptons too!) can decay in several distinct ways; a few examples for illustration purposes are [293]:

Decay results		**Decay results**			
$\Lambda^0 \to p^+ + \pi^-$	$(35.8 \pm 0.5)\%$,	$\to n^0 + \pi^0$	$(63.9 \pm 0.5)\%$,	etc.	
$K^+ \to \mu^+ + \nu_\mu$	$(63.44 \pm 0.14)\%$,	$\to \pi^+ + \pi^0$	$(20.9 \pm 0.12)\%$,		(2.62)
$\to 2\pi^+ + \pi^-$	$(5.590 \pm 0.031)\%$,	$\to \pi^+ + e^+ + \nu_e$	$(4.98 \pm 0.07)\%$,	etc.	
$\tau^- \to \mu^- + \nu_\tau + \overline{\nu}_\mu$	$(17.36 \pm 0.05)\%$,	$\to$ all else:	$(82.64 \pm 0.05)\%$.		

---

[30] Roughly, all *laws of Nature*, as they are understood here, contain both conservation laws and interaction laws.

One of the tasks of elementary particle physics is the computation of the relative probabilities of decay, as well as the total lifetimes of various particles. This includes both the elementary particles [☞ Table 2.3 on p. 67], such as the τ-lepton, and also the bound states of these elementary particles. This second group, much larger and growing, consists of hadrons (mesons and baryons) as well as the so far only hypothetical and experimentally unverified bound states that consist purely of gluons,[31] and possibly also the so-called exotic hadrons – quark bound states that are neither mesons, $(q\,\bar{q})$, nor baryons, $(q\,q\,q)$; for example, so-called *dibaryons* are hypothetical bound states of six quarks.

> ***Conclusion 2.4*** *The primary focus of the so-called "elementary particle physics" is on the elementary particles as identified in the Standard Model [☞ Table 2.3 on p. 67]. However, this then covers also the dynamics of these elementary particles, and so also their bound states: all mesons and all baryons [☞ Table 2.5 and Section 11.2].*

By the feature indicated in this conclusion, high energy physics currently differs from all other disciplines in physics: The domains of study of several closely related physics disciplines are sketched in Table 2.5. Unlike all other disciplines in this table, "elementary particle physics" (also known as "high energy physics") studies (at least) *two* levels of elementarity.

Table 2.5 The domains of several physics disciplines of "small" systems and objects

Discipline	Domain of study
Molecular physics	Molecules (chemically bound states of atoms)
Atomic physics	Atoms (electromagnetic bound states of a nucleus and electrons)
Nuclear physics	Atomic nuclei (bound states of protons and neutrons)
Elementary particle physics	Elementary particles [☞ Table 2.3 on p. 67]
a.k.a. High energy physics	Bound states of these (mesons and baryons)

Finally, by the end of the twentieth century, high energy physics had also brought on an essential shift in the understanding of the Democritean idea of "elementary *particles*": The hierarchy

1. molecules consist of atoms,
2. atoms consist of electrons and a nucleus,
3. nuclei consist of nucleons (protons and neutrons),
4. nucleons (and all other hadrons) consist of quarks,

experimentally stops here, for now. It is reasonable to expect that contemporary "high energy physics" will soon effectively split into "hadron physics" and "fundamental physics," although their respective domains do not yet seem to be sufficiently differentiated.

According to (super)string theory, this hierarchical halt is also conceptually significant: In that theoretical system, the fundamental objects are not any new (smaller, constituent) *particles*, but (super)strings; the particles of the Standard Model, as given in Table 2.3 on p. 67, are not bound states of these more elementary (super)strings, but are their *modes of vibration*. This is an essential shift in understanding: quarks, leptons, gauge and Higgs bosons are not at all bound states consisting of other, more elementary things! Rather, the same string contains amongst its vibrations (its Fourier modes) all the elementary particles of the Standard Model (and indefinitely more) *simultaneously* [☞ Chapter 11].

---

[31] Unlike the chargeless photons in abelian (commutative) gauge theory of quantum electrodynamics, the mediators of non-abelian (non-commutative) interactions in quantum chromodynamics (the gluons) themselves have color charge and so can form bound states, so-called *glueballs* [☞ Chapter 6.2].

Returning then to elementary *particle* physics, consider the correlation between the mass of a fundamental interaction mediator and the range of that interaction. In 1931, Yukawa reasoned that the total energy is $E = c\sqrt{m^2c^2 + \vec{p}^{\,2}} \geqslant mc^2$ for a mediating particle of mass m. To produce such a particle during the interaction, at least mc^2 energy is needed. Heisenberg's indeterminacy relations permit "borrowing" that much energy for no longer than $\sim \frac{\hbar}{E} \leqslant \frac{\hbar}{mc^2}$, during which this mediating particle may traverse a distance no larger than *[✎ why?]*

$$R \sim \frac{\hbar}{mc}. \tag{2.63}$$

Numerical factors such as $\frac{1}{2}$ in $\triangle E \triangle \tau \geqslant \frac{1}{2}\hbar$ were neglected here as the estimate (2.63) is a rough *upper limit*. Using the fact that the strong interaction must have a range that is *at least* comparable to the size of the atomic nucleus – so as to keep the nucleons in the bound state – Yukawa estimated the mass of the mediators for the strong nuclear interaction as $m \sim \frac{\hbar}{Rc} \approx 200\,\text{MeV}/c^2$.

Digression 2.6 Warning! Using the same reasoning to estimate the range of the electromagnetic interaction from the size of the atom, $a_0 = 5.291\,772\,108 \times 10^{-10}$ m, implies that the photon mass is $\sim$4 keV$/c^2$ – which is wrong. Of course, we know that the range of the electromagnetic interaction is infinite, which agrees with the relation (2.63) and the fact that the photon mass is zero. The error in the first estimate stems from the fact that the binding energy of the hydrogen atom (1.31) is less than the rest energy of the electron by the dimensionless factor $\alpha^2 \sim 5.33 \times 10^{-5}$; bound states where the binding energy is a few orders of magnitude smaller than the rest energy are called "weakly bound" [☞ relations (1.40)]. Heisenberg's relations are *inequalities*, and thus can only provide a *lower limit* for the mass of the mediating particle from the size of the bound state, and usefully so only for "strongly bound" states!

In the case of weak interactions, the mass estimate for the mediating bosons, $W^\pm, Z^0$, is additionally hampered by the fact that there do not exist states bound by the weak nuclear interaction. Since the β-decay evidently does happen within the atomic nucleus, we know that the range may well be of the order of the atomic nucleus diameter, but this does not permit estimating either limit for the range: The range may be smaller and the β-decay occurs at distances much smaller than the nucleus diameter. Or, it may be much larger than the nucleus – the involved particles are confined within the nucleus anyway, by the strong interactions.

Complementary to the range, we may compare the times required for a decay to happen. Strong interaction decays typically happen within 10^{-23} s, while electromagnetic decays occur within 10^{-16} s – ten million times slower. Weak interactions, however, vary: decays may be as fast as 10^{-13} s (for the τ-lepton) to as long as 881.5 s (for the free neutron) [293] – which is a spectrum of 16 orders of magnitude! If the decay results in photons, it happened by means of the electromagnetic interaction; similarly, the appearance of a neutrino in the decay results is the "hallmark" of weak interactions. For decays where the result contains neither a photon nor a neutrino, the type of the interaction is harder to determine, so the duration of the decay is a useful indicator.

While it is known that all particles except e^-, p^+, ν_e and γ decay, the decay patterns of all particles are very regular. A systematic analysis of their decays, and also of their inelastic collisions and scatterings then implies:

Conclusion 2.5 *All particles decay into lighter particles, and in all manners that are permitted by conservation laws. (The electron, for example, does not decay as there are no lighter particles into which to decay.)*

This conclusion is related to Conclusion 2.3 on p. 56. This logic led to a successful application of the $SU(3)_f$ and later also of the $SU(4)_f$ approximate symmetries, as well as several conservation laws that became an integral part of the Standard Model. The subsequent review of these laws will have to be amended later, when the technical details become familiar.

2.4.2 Strict conservation laws

By "strict conservation laws" we understand those laws that hold for all interactions and in all situations. For each of these laws, the Standard Model exhibits an explicit symmetry, linked with that conservation law by way of Emmy Noether's theorem. The book [383] is dedicated to various forms and applications (not only in physics!) of this important theorem.

4-momentum, angular momentum, parity

One of the most important lessons from the historical review in Section 2.3 is the reliance on conservation laws of the 4-momentum: both the (relativistic) total energy and the vector of linear momentum are strictly conserved quantities in so-called real states – i.e., in states where these quantities can be observed and measured. On the other hand, in *virtual* states neither the total energy nor the momentum can be measured, and the conservation laws are not applicable. Thus, it is not that the energy and/or momentum conservation law is violated in virtual states, but rather there is neither measurable energy nor measurable momentum for which to apply the law.

The conservation of the 4-momentum is the consequence of Noether's theorem for the spacetime translation symmetry, for a system's real states.

Digression 2.7 Heisenberg's indeterminacy relations are not infrequently cited as the assertion that physical quantities that can be simultaneously measured must correspond to operators that commute. That, in fact, is not quite true, because it neglects the essential dependence in quantum mechanics on the state, or a class of states, in which the considered system is prepared.

A very general proof[32] of this statement follows from considering two Hermitian operators, A and B, which define

$$C := -i\big[A\,,\,B\big], \qquad C^\dagger = C. \tag{2.64a}$$

Defining $A_0 := A - \langle A\rangle$ and $B_0 := B - \langle B\rangle$, we have

$$\big[A_0\,,\,B_0\big] \;:=\; \big[(A - \langle A\rangle)\,,\,(B - \langle B\rangle)\big] \;=\; iC, \tag{2.64b}$$

so that

$$0 \leqslant \langle\big|A_0 - i\omega B_0\big|^2\rangle = \langle A_0^2\rangle - i\omega\langle\big[A_0\,,\,B_0\big]\rangle + \omega^2\langle A_0^2\rangle \tag{2.64c}$$

$$= \Delta_A^2 + \omega\langle C\rangle + \omega^2\Delta_B^2. \tag{2.64d}$$

The right-hand side expression is minimized by $\min(\omega) = -\langle C\rangle/2\Delta_B^2$, producing

$$\Delta_A^2\,\Delta_B^2 \;\geqslant\; \tfrac{1}{4}\langle C\rangle^2, \qquad \text{that is,} \qquad \Delta_A\,\Delta_B \;\geqslant\; \tfrac{1}{2}\,|\langle C\rangle| = \tfrac{1}{2}\,|\langle[A,B]\rangle|\,, \tag{2.64e}$$

which are Heisenberg's indeterminacy relation for the physical quantities represented by the operators A, B. This manifestly depends on the state in which the indicated expectation values $\langle C\rangle$ and $\langle[A,B]\rangle$ are computed – and may well be zero although $[A,B] \neq 0$.

---

[32] This follows the variational derivation in Refs. [295, 97], refining the original derivation by Robertson [460] and Schrödinger [476]; see also Ref. [242].

Example 2.1 In the best known example, we have $A = p_x$ and $B = x$, and $C = \hbar\mathbb{1}$ is a constant, and the *non-vanishing* of the right-hand side of equation (2.64e) is in fact state independent. However, for the case of the angular momentum,

$$\big[\,J_x\,,\,J_y\,\big] = i\hbar\,J_z \qquad\Rightarrow\qquad \Delta_{J_x}\,\Delta_{J_y} \geqslant \tfrac{1}{2}\big|\langle J_z\rangle\big|. \tag{2.65}$$

For states where $\langle J_z\rangle = 0$ we have that $\Delta_{J_x}\,\Delta_{J_y} \geqslant 0$. Thus, although the operators J_x and J_y do not commute, the product of their indeterminacies may well vanish in states of the system where $\langle J_z\rangle = 0$. In those quantum states, Heisenberg's indeterminacy principle does not preclude the simultaneous measurement of J_x and J_y although they do not commute as operators.

Just like 4-momentum, angular momentum is strictly conserved in real states. The conservation law is a consequence of Noether's theorem for rotation symmetry.

Parity, P, is the operation that changes the sign of all Cartesian spatial coordinates; in spherical coordinates, this is the $(r,\theta,\phi) \to (r,\pi-\theta,\phi+\pi)$ transformation. With respect to this, so-called polar vectors (position, velocity, acceleration, electric field, etc.) all change sign. In turn, so-called axial vectors (angular momentum, torque, magnetic field, etc.) do not change sign. Scalar functions of position (temperature, atmospheric pressure, density, etc.) are invariant and do not change sign, while *pseudo-scalar* functions of position (e.g., the volume element $\mathrm{d}^3\vec{r}$) change sign. Tsung Dao Lee and Chen Ning Yang discovered that parity is strictly conserved in all electromagnetic and strong processes, but that by 1956 parity conservation had not been verified in weak interactions. Thus, they proposed several direct experimental tests. During 1956–7, Madam Chien-Shiung Wu found, with her collaborators, clear indication of P-violation in the ${}^{60}_{27}$Co β-decay, which was immediately confirmed by R. L. Garwin, L. Lederman and R. Weinrich by means of precise measurements of cascading decays (2.21). It later turned up that R. T. Cox, C. G. McIlwraith and B. Kurrelmeyer had published experimental results back in 1928 [118, 245] about double scattering β-particles ($e^{\pm}$), which indicated P-violation, but those 28 years earlier no one was willing to consider that as an explanation.

Electric charge

The Maxwell equations (5.72) straightforwardly produce the so-called equation of continuity:

$$\begin{array}{rcl} \vec{\nabla}\cdot\vec{E} = \dfrac{1}{4\pi\epsilon_0}4\pi\,\rho_e & \Rightarrow & \dfrac{\partial(\vec{\nabla}\cdot\vec{E})}{\partial t} = \dfrac{1}{4\pi\epsilon_0}4\pi\,\dfrac{\partial\rho_e}{\partial t},\\[2ex] \vec{\nabla}\times(c\vec{B}) - \dfrac{1}{c}\dfrac{\partial\vec{E}}{\partial t} = \dfrac{1}{4\pi\epsilon_0}\dfrac{4\pi}{c}\vec{\jmath}_e & \Rightarrow & \underline{\vec{\nabla}\cdot\big(\vec{\nabla}\times(c^2\vec{B})\big) - \dfrac{\partial(\vec{\nabla}\cdot\vec{E})}{\partial t} = \dfrac{1}{4\pi\epsilon_0}4\pi\vec{\nabla}\cdot\vec{\jmath}_e,}\\[2ex] & & \Rightarrow\quad 0 = \dfrac{\partial\rho_e}{\partial t} + \vec{\nabla}\cdot\vec{\jmath}_e, \end{array} \tag{2.66}$$

since $\vec{\nabla}\cdot(\vec{\nabla}\times\vec{X}) \equiv 0$ for all $\vec{X}$. It follows, integrating

$$\frac{\partial\rho_e}{\partial t} = -\vec{\nabla}\cdot\vec{\jmath}_e \qquad\Rightarrow\qquad \frac{\mathrm{d}Q_{e,V}}{\mathrm{d}t} = -\oint_{\partial V}\mathrm{d}^2\vec{r}\cdot\vec{\jmath}_e, \tag{2.67}$$

so that the change of the total amount of electric charge $Q_{e,V}$ contained within a volume V equals the flux of the electric current through the (surface) boundary of that volume. The conservation law of electric charges is thus an exact and *inevitable* consequence of the fundamental laws of electromagnetism.

A violation of this law in any process would then indicate a contradiction with the Maxwell equations, and so also with electrodynamics as a whole. In the weak interactions, such as

$$\begin{tikzcd} & e^+ + W^- + p^+ \arrow[dr] & \\ \overline{\nu}_e + p^+ \arrow[ur] \arrow[dr] & & e^+ + n^0, \\ & \overline{\nu}_e + W^+ + n^0 \arrow[ur] & \end{tikzcd} \tag{2.56}$$

the electric charge of individual particles transfers: the neutral antineutrino becomes the positive positron and the positive proton becomes the neutral neutron. However, the *total* electric charge remains conserved in any arbitrarily small volume that contains the interacting particles. In the alternative intermediate processes, we see that the $W^\pm$ also carry electric charge, so that the electric charge is conserved in each of the individual elementary processes:

$$\overline{\nu}_e \to e^+ + W^-, \qquad u \to W^+ + d, \qquad W^- + u \to d, \qquad W^+ + \overline{\nu}_e \to e^+. \tag{2.68}$$

Digression 2.8 The very fact that the weak nuclear interaction mediators carry also electric charge indicates that the weak nuclear interaction is not fully independent of the electromagnetic one. Chapter 7.2 will more precisely examine this link. However, one of the *consequences* of this link has already emerged, in the generalized GNN formula (2.44b).

Color

Chapter 6.1 will show that the *color* in quantum chromodynamics generalizes the electric charge: whereas the electric charge is a scalar quantity, color is a 3-component quantity, i.e., a 3-vector in an abstract 3-dimensional space, just as the 3-vectors of position, velocity and force are vectors in the "real" space in which we ourselves move.

The fundamental differential equations of chromodynamics are the corresponding generalization of the Maxwell equations, and so also follow a corresponding conservation law of *color* as a charge. During elementary chromodynamical processes, quarks change their color, but this change is carried by gluons (the strong nuclear interaction mediators), so that color is conserved in every process. Since all detectable (real) particles are *colorless*, the color conservation law and the corresponding global symmetry are somewhat trivial. However, the gauge (local, i.e., spacetime variable) color symmetry is the reason for the existence of the strong nuclear interaction; see Chapter 6.

Lepton numbers

Unlike electric charge and color, which are conserved charges of gauge symmetries and which thus correspond to electromagnetic and strong nuclear interactions, the lepton numbers are strictly conserved, but are not the conserved charges of a gauge symmetry and do not correspond to any interaction. For example, in the decay

$$\mu^- \to e^- + \overline{\nu}_e + \nu_\mu, \tag{2.69}$$

the muon lepton number ($L_\mu = +1$) is carried by μ^- and ν_μ, and we have $L_\mu = +1$ input, $L_\mu = +1$ output. The electron lepton number is carried by e^- and ν_e: $L_e(e^-) = L_e(\nu_e) = +1$, so that $L_e(\overline{\nu}_e) = -1$. Here we have $L_e = 0$ input, $L_e = +1 + (-1) = 0$ output. Following the original proposal [319], a systematic analysis of all so far observed processes (except for neutrino mixing; see Section 7.3.2) indicates a strict conservation of all three lepton numbers, L_e, L_μ, L_τ, as defined in the table:

	ν_e, e^-	$\overline{\nu}_e, e^+$	ν_μ, μ^-	$\overline{\nu}_\mu, \mu^+$	ν_τ, τ^-	$\overline{\nu}_\tau, \tau^+$
$L_e =$	$+1$	-1	0	0	0	0
$L_\mu =$	0	0	$+1$	-1	0	0
$L_\tau =$	0	0	0	0	$+1$	-1

(2.70)

An analogous conservation of quark numbers, *separately* for the (u, d), (c, s), and (t, b) pairs does not exist, because of the CKM mixing (2.53) of so-called "lower" quarks, d, s, b. The question of lepton mixing, i.e., neutrino mixing, will be addressed in Section 7.3.2; let us note here merely that this possibility was proposed back in 1962 [353], although there was no strong experimental indication until recently that such a mixing really happens [369, 370].

In this sense is the reason for the existence of the (approximate) conservation law of three separate lepton numbers and the absence of a conservation law of three separate quark numbers a phenomenological and not a fundamental law – and an open question☝!

Baryon/quark number

The quark model redefined the baryon number simply as the triple of the quark number, where antiquarks have negative quark number. In the Standard Model, that definition remains, and also explains the absence of a meson conservation number: since mesons are $(q\overline{q})$ bound states, their quark number is zero. Since quarks cannot be extracted, it remains a convention to count baryons, and quarks have $\frac{1}{3}$ of the baryon number.

The baryon number conservation law is also strict – in that it holds in all processes. However, just as the (separate) lepton number conservation laws, this too is a phenomenological and not a fundamental law.

2.4.3 Approximate conservation laws

Besides strict conservation laws, there also exist approximate conservation laws, which are nevertheless useful precisely because of their approximate validity, whereby they help in estimates and computations.

Flavor

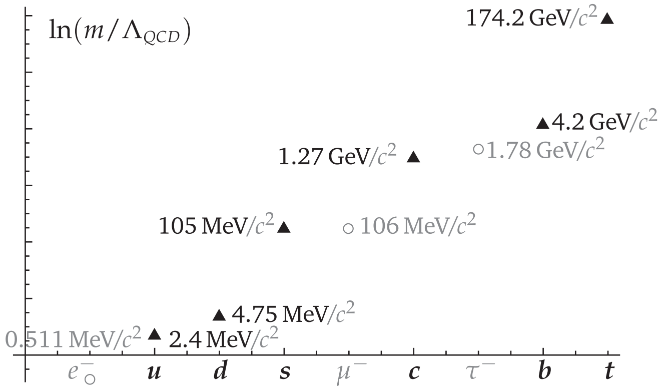


Figure 2.1 Quark (▲) and charged lepton (○) masses plotted on a logarithmic scale.

Table (2.44a) shows that the differences between the consecutive quark masses grow with these masses, as seen on the plot 2.1. In experiments done at the average energy of $\Lambda_{QCD} = 200\,\text{MeV}/c^2$ per process and with the experimental error at about 10% (so about $20\,\text{MeV}/c^2$), it is not possible to distinguish u- and d-quarks purely by their masses; within experimental error, their masses are the same. On the other hand, there is enough energy to produce an s-quark, which indeed can be distinguished from the u- and the d-quarks purely by its mass: $105 \pm 20\,\text{MeV}/c^2$ cannot be confused for m_u, m_d even when identified within the $\pm 20\,\text{MeV}/c^2$ experimental error.

Nevertheless, Gell-Mann proposed to:

1. First consider m_u, m_d, m_s as sufficiently near in masses to be distinguished; this indicated an $SU(3)_f$ symmetry.
2. Then take into account the difference between m_s *vs.* $m_u \approx m_d$; this breaks the symmetry $SU(3)_f \to SU(2)_{u,d}$.

This strategy led to his classification system "eightfold way," the plots such as (2.31), (2.32) and (2.35), the phenomenological formulae (2.33)–(2.34), and also to the discovery of the Ω^- baryon. Thereby, Gell-Mann introduced and established the use of symmetry – even if only approximate.

Within the Standard Model, such classifying schemes are, based on the "flavor" $SU(n)_f$ symmetry, very clearly phenomenological schemes. The conservation laws of individual "flavors" or groups of "flavors" are also only approximate rules, broken by the CKM mixing (2.53)–(2.55).

Digression 2.9 This induced the idea that approximate symmetries are (perhaps, sometimes) "only" broken symmetries, prompting us to uncover the reason and mechanism of breaking. In this sense is the origin of quark masses, and lepton masses too, as well as the CKM matrix, one of the basic questions to which the Standard Model has no answer☝. The quest for this origin is one of the basic motivations for most proposals that go "beyond" the Standard Model. This includes various electroweak and strong interaction unification models, and in these models at least some of the unexplained characteristics of the Standard Model are supposed to be derived and "predicted."

The OZI rule

There is a very general regularity in decays: The speed and probability of a decay,

$$X \longrightarrow Y_1 + Y_2 + \cdots + Y_k, \tag{2.71}$$

both grow with the change in mass, $\triangle m := (m_X - \sum_i m_{Y_i})$. Thus, between two decays that occur by means of the same kind of interaction, the one for which $\triangle m$ is larger happens more often. Deviations from this regularity require an explanation.

In the 1960s, Susumu Okubo [392], George Zweig [593] and Jugoro Iizuka [289] independently discovered a significant correlation: decays that require the full annihilation of all "incoming" quarks and antiquarks are delayed (i.e., the probability of such decays is diminished) as compared to decays of the same system where at least some of the incoming quarks or antiquarks pass through into the decay result. During the 1960s, his correlation so successfully "explained" the delayed decays that it acquired the nickname the "OZI rule." For example, the probability of the $\phi \to \pi^+ + \pi^- + \pi^0$ decay is diminished as compared to the probability of the $\phi \to K^+ + K^-,\ K^0 + \overline{K}^0$ decay; experiments show that ϕ decays over 83% of the time into kaons and not into pions, although

$$\triangle m(\phi \to 2K) \approx 32.1\ \text{MeV}/c^2, \qquad \text{while} \qquad \triangle m(\phi \to 3\pi) \approx 605\ \text{MeV}/c^2. \tag{2.72}$$

Analogously, the J/ψ particle was supposed to decay predominantly into the pair of mesons $D^{+}=(c\,\overline{d})$ and $D^{-}=(d\,\overline{c})$. However, the total mass of the $D^{+}+D^{-}$ meson pair is *larger* than the mass of the J/ψ meson, so that the decay $J/\psi\to D^{+}+D^{-}$ is *kinematically forbidden* [☞ Section 3.2], and only the decays into charmless hadrons remain possible, for all of which the probability is diminished by the OZI rule. This then is the combination of reasons that induces J/ψ to have unusually long ($\sim 10^{-20}$ s) lifetime, which is some 1,000 times longer than most other hadrons.

Symmetries and models

Symmetries have now been mentioned and even used several times, relying on the intuitive understanding of their nature and physical meaning. However, to be more precise – and especially with the discussion of (mathematical) models started in Section 1.1.2 in mind – recall that these mathematical models serve to faithfully reproduce all characteristics of the considered system. The model is therefore automatically identified with the physical system, object or quantity that the model represents [☞ Section A.1.3].

For example, strictly speaking, the 3-dimensional vector $\vec{B}$ is the abstract mathematical construct used as a model for the magnetic field of a concrete magnet, in a concrete point of space and in a concrete moment in time. The union of continuously many vectors $\vec{B}$ in space around the particular magnet in the same moment in time provides the abstract mathematical construct (it is impossible to measure continuously many points) that is automatically identified with the concrete magnetic field of that magnet. The abstract mathematical property of this union of vectors $\vec{B}$ – that it does not change if the whole union is rotated about the axis of the magnet itself – is automatically ascribed to the concrete magnetic field of the concrete magnet.

By the same token, we have more generally:

> ***Conclusion 2.6*** *Symmetries and other significant properties of the abstract mathematical model are automatically ascribed to the concrete physical system, object or quantity that the model faithfully represents.*

We then say that the (concrete, physical) magnetic field of the magnet has axial symmetry, even though this symmetry – strictly speaking – is a property of the mathematical model of this magnetic field. It is therefore of the essence that models do not introduce unnecessary (fictitious) degrees of freedom, concepts and properties.

> ***Conclusion 2.7*** *Ideally, the mathematical models of physical systems, objects and quantities must be* ***optimal****: minimal in the number and complexity of structure of intermediary and auxiliary means, (self-)consistent and faithful in representing the physical system, object and quantity for the description of which it is used.*

On one hand, this requirement of optimality reminds us that physics does not describe Nature directly, but *through* mathematical models that are continually improved. On the other hand, this requirement is a variation of Ockham's principle, which crucially limits the possibilities at our disposal when improving the existing models or creating new ones.

This practice largely determines the development of physics.

2.4.4 Exercises for Section 2.4

✎ ***2.4.1*** *A particle for which the relation (2.1) determines the speed and the ratio of electric charge by mass enters under the same conditions into a region where, however, now the* *magnetic field is turned off. Compute the distance and direction of deflection in the (y,z)-plane from the x-axis, when the particle has traversed the length ℓ along the positive x-axis, and show that this deflection again depends on the ratio of charge by mass.*

That is, show that successive measurements on the **same** *cathode ray, with either one, or the other, or both fields $\vec{E} = E_0\,\hat{e}_y$ and $\vec{B} = B_0\,\hat{e}_z$ cannot determine independently both the electric charge and the mass of the particles that make up the cathode ray.*

✎ ***2.4.2*** *A photon of energy $E_\gamma = h\nu$ and linear momentum $\vec{p}_\gamma = \frac{1}{c}E_\gamma\,\hat{e}_x$ collides with an electron at rest. Upon the collision, the photon continues in the direction $\cos(\phi)\,\hat{e}_x + \sin(\phi)\,\hat{e}_y$ with the energy E'_γ and linear momentum $\vec{p}'_\gamma$, while the electron recoils in the direction $\cos(\theta)\,\hat{e}_x - \sin(\theta)\,\hat{e}_y$ with the linear momentum $\vec{p}_e$. Show that the linear momentum and energy conservation laws (a) produce the result (2.5), and (b) forbid that a free electron simply absorbs a photon.*

✎ ***2.4.3*** *Show that the energy and linear momentum conservation forbids* **all** *processes that may be obtained from (2.15)–(2.16) when either of the two photons is deleted and the remaining particles are rearranged using the crossing symmetry and the principle of detailed balance.*