Scientific Background on the Nobel Prize in Physics 2008

Broken Symmetries

compiled by the Class for Physics of the Royal Swedish Academy of Sciences
Symmetries and Groups

Mankind has been interested in symmetries in Nature as long as history can tell. Symmetric objects are so distinguished in the world around us that they have often been given special status. The obsession of the Greeks with symmetries led them to classify many noteworthy shapes, and many cultures have used symmetries and symmetric objects as symbols in their lives. Of course, most shapes in Nature display little or no symmetry, but many are almost symmetric. An orange is almost a perfect sphere, while the human being is almost symmetric about an axis through the body, but not quite. There must be reasons for this. The symmetric shape of the orange helps it to grow, but the small imperfection helps it, too. There are also perfectly symmetric situations in life where we must break the symmetry to proceed. At a round dinner table with napkins placed between the plates, should one take the one to the left or the one to the right? Both situations represent identical results. Someone has to start, however, by breaking the symmetry. Once this is done, everyone knows which side to choose. We shall see that this phenomenon occurs at the fundamental levels of physics, too. It took the mathematicians up to the 19th century to systematically study the mathematical structure behind symmetric patterns, and in the 20th century the mathematical subject Group Theory became one of the most important tools for understanding modern physics.

It was known in the 19th century that Newtonian mechanics is invariant under rotations and translations. Maxwell’s equations are, however, invariant under a different set of transformations found by Hendrik Lorentz (Nobel Prize 1902) [1]. It took the brilliance of the French mathematician and physicist Henri Poincaré [2] to realize in 1905 that these symmetry transformations indeed form a group and hence were amenable to a formidable machinery of mathematics. This is, however, historically shadowed by Albert Einstein’s (Nobel Prize 1921) [3] discovery of the theory of Special Relativity, also in 1905, in which Lorentz transformations also underlie mechanics. When quantum mechanics was discovered in the 1920s, it looked at first as if non-relativistic quantum mechanics would be enough to explain the spectrum of the hydrogen atom. The discovery of the spin of the electron, however, implied a finer structure in the spectrum, which prompted Wolfgang Pauli [4] (Nobel Prize 1945) to add spin corrections to the Schrödinger equation [5] (Nobel Prize 1933 to Erwin Schrödinger). Then Paul Dirac gave a completely different and better explanation in 1928 [6] (Nobel Prize 1933), to the astonishment of the physics community, when he introduced his equation, the Dirac equation, which satisfies invariance under Special Relativity. Dirac used his knowledge of group theory and found a specific representation of the Poincaré symmetry. His equation then included the spin naturally. A complete treatment of this symmetry was finally given in 1939 by Eugene Wigner [7] (Nobel Prize 1963), who used full mathematical formalism to treat the Poincaré symmetry and concluded that the representations can be described by an integer or half-integer spin (in Planck’s constant) and a given mass. (Wigner had been encouraged to carry out this analysis by Dirac as early as 1927. The paper was finished in 1935 but it was several years before it was published.) The Poincaré group is an example of a dynamical symmetry, i.e. one in which dynamical variables enter.

The Poincaré group is also an example of an infinite group, one in which there are infinitely many possible symmetry operations. In the 1920s, finite symmetry transformations involving space and time that could be adjoined to the Poincaré symmetries were also discovered. In 1924, Otto Laporte [8] found that energy levels in complex atoms could be classified according to even or odd levels. Laporte found that when a photon is emitted, the level changes from an even to an odd or vice versa. In 1927, Wigner [9] took a profound step to prove that the empirical rule of Laporte is a consequence of the reflection symmetry of the electromagnetic interactions. (This means that a process seen in a mirror is as likely as the
Hence, this symmetry involves the parity operation in which the space coordinates are reflected. The symmetry group involves two elements, the identity and the parity operation $P$. The elementary particles must also be given a definite parity. (The parity of the photon is negative, hence the emission from an even to an odd or the opposite.) Dirac also found another finite symmetry operation, when he realized that his equation in fact describes two spin-$1/2$ particles with opposite charge. He first thought the two were the electron and the proton, but it was then pointed out to him by Igor Tamm and Robert Oppenheimer that they must have the same mass, and the new particle became the anti-electron, the positron. It was discovered by Carl Anderson [10] in 1932 (Nobel Prize 1936). The charge conjugation symmetry $C$ changes a particle to an antiparticle. Wigner introduced a third finite symmetry in 1932 [11], the time reversal symmetry $T$. This symmetry operation consists of changing time to its negative. If the full operation $CPT$ is acted on the Dirac equation, its effect is unity. It was later found by Julian Schwinger [12] (Nobel Prize 1965) in 1951, Gerhard Lüders and Pauli [13] and John Bell [14] in 1954 that under quite general conditions $CPT$ is conserved in a relativistic quantum field theory.

Physicists are most often satisfied to consider only infinitesimal transformations of the infinite groups. These satisfy an algebra and most often the Poincaré algebra is used instead of the group. It has ten elements, four translations, three rotations and three boosts, which satisfy certain commutation relations (In physics it is called $SO(3,1)$). Physicists often do not distinguish between the group and the algebra with the notation.). This is a so-called Lie algebra, and many other such algebraic structures have found their way into physics. The first one was the isospin symmetry ($SU(2)$). It was introduced by Werner Heisenberg [15] (Nobel Prize 1932) to explain symmetries of the then newly discovered neutron and the proton. This symmetry assumes that the proton and the neutron have the same mass and are indistinguishable under the strong interactions. Non-dynamical symmetries like this one are called internal symmetries. They transform in an internal space.

All these symmetries are global symmetries. They act the same way at each space-time point. The notion of a local symmetry, a gauge symmetry, i.e. a symmetry in which the transformations vary from point to point in space-time, was introduced by Hermann Weyl. In 1918 he discussed such a notion in the context of gravity and in 1929 [16] he realized that electromagnetism can be understood as a realization of such a gauge invariance. By letting the theory be invariant under local phase rotations of the electron wave function, the full theory follows and the electromagnetic field is introduced as a gauge field. The gauge group is $U(1)$, which is an abelian group, meaning that transformations commute with each other. Hence electromagnetism is called an Abelian Gauge Theory.

The four fundamental interactions

After the neutron had been discovered by James Chadwick in 1932 [17] (Nobel Prize 1935) it became clear that there are specific forces acting within the atomic nucleus, and in 1934 Wigner showed that there must be two different nuclear forces at play, a weak one that is responsible for radioactivity and a strong one that binds the protons and the neutrons together. There are then four basic forces in Nature, the strong and the weak nuclear force, the electromagnetic force and gravity. Much effort in particle physics since the 1930s has been made to understand these four forces.

The first force to be understood in a quantum mechanically consistent way was Quantum Electrodynamics (QED). The key persons in this development were Richard Feynman [18], Julian Schwinger [19] and Sin-Itiro Tomonaga [20] (who shared the Nobel Prize 1965) and
Freeman Dyson [21]. They also found that a seemingly consistent systematic perturbation theory could be formulated for this theory. They also found that all infinities that appear at each order in the perturbation expansion could be absorbed into a small set of parameters that could then be defined by their experimental values; the theory is then called renormalizable. Gauge invariance played a vital role here. A very important fact that follows from gauge invariance is that the theory must have a conserved charge, in this case the electric charge. By using Emmy Noether’s analysis of symmetries [22] it is found that an infinite symmetry leads to a conserved current, which in its turn leads to a conserved charge. In QED the current is with \( e(x) \) the electron field

\[
J^\mu_e = e(x)\gamma^\mu e(x).
\]

The conservation condition is then

\[
\partial_\mu J^\mu_e = 0,
\]

where \( \partial_\mu = \partial / \partial x^\mu \). The Hamiltonian in QED can now be written as

\[
H = eJ^\mu_e A_\mu,
\]

where \( e \) (the electric charge) is the coupling constant in QED. If the electromagnetic field couples to other fields, the current will consist of a sum of terms like the one above, and all parts of the current of a given charge will couple with the same strength \( e \). This coupling is universal. In the renormalization, it is the coupling constant that is renormalized. If a universal coupling is found for different particles, it can be interpreted as a sign of an underlying renormalizable theory with an infinite symmetry leading to a conserved current.

Since QED paved the way for describing relativistic many-body theories, it was natural to attempt a similar approach for the theory proposed for the strong interactions by Hideki Yukawa [23] in 1935 (Nobel Prize 1949). He had proposed that the interactions between the nucleons are mediated by pions with a small but non-zero, mass leading to a short range interaction. This theory was also found to be renormalizable, but the coupling strength was found to be larger than 1, so a perturbation expansion is meaningless. The failure of the Yukawa theory led the physics community to doubt the relevance of relativistic quantum field theory, and many attempts to find alternative formalisms for the strong force were made in the 1950s.

**Parity Non-Conservation and the Weak Interactions.**

The groundwork for a field theoretic formulation of the weak interactions was carried out by Enrico Fermi [24] (Nobel Prize 1938) in 1934. The observation of a continuous electron spectrum in the \( \beta \)-decay of the nucleons had convinced Pauli [25] that the fundamental process is

\[
_z X \rightarrow_z X + e^- + \nu,
\]

where \( \nu \) is the neutrino postulated by Pauli. Fermi then suggested an interacting Hamiltonian of the form

\[
H = -G \pi(x)\gamma^\mu p(x)\bar{\nu}(x)\gamma_\mu \nu(x) + h.c.
\]
This form can be seen as a current-current interaction with a point-like interaction. During the next two decades, many new experiments were performed and new particles and new decays were discovered. To bring order among the new particles, Murray Gell-Mann [26] (Nobel Prize 1969) and Kazuhiko Nishijima [27] introduced a new quantum number, \textit{strangeness}. This new quantum number explained why certain particles are always produced in pairs in scatterings where the initial states have no strangeness; the particles in the pair have opposite strangeness. A big problem remained. There were two particles called 0 and \(\tau\), which apparently had the same quantum numbers and mass but decayed into two and three pions respectively.

By checking the parity of the final state, it was found that they have different parity. Could there be two particles with the same properties but with different parity? Chen-Ning Yang and Tsung-Dao Lee [28] solved this problem in 1956 (Nobel Prize 1957). They carefully investigated the experiments that had been performed and found that the believed parity laws had not been tested in weak processes. For the strong and electromagnetic interactions, on the other hand, there was convincing evidence to support the parity law. They then proposed to test it in the \(\beta\)-decay of \(^{60}\text{Co}\) and the muon respectively. These experiments showed unambiguously that the weak interaction is not invariant under a mirror transformation in space [29], and the problem of the 0 and \(\tau\)-particles was solved. They are simply the same particle, the \(K\)-meson or the kaon, whose weak decay does not respect parity. It is interesting to note that Dirac, in a contribution to Einstein’s 70th birthday in 1949 [30], wrote that he saw no reason why \(P\) and \(T\) invariance should be conserved, but that the existing theories apparently did satisfy this requirement.

Both Abdus Salam [31] (Nobel Prize 1979) and Lev Landau [32] (Nobel Prize 1962) argued shortly after the discovery of non-conserved parity in the weak interactions that the interaction can be invariant under chiral transformations of the neutrino. It can be “left-handed” and exist only with one kind of helicity. It can then be described by a two-component spinor, a Weyl spinor. It was now possible to try to extend the current-current interaction of Fermi, which is of vector type, to any of the five Lorentz covariant currents \(S\) (scalar), \(P\) (pseudoscalar) \(V\) (vector), \(A\) (axial vector) or \(T\) (tensor). Experiments gave different answers and did not provide a definite clue. Robert Marshak and E.C.G. Sudarshan [33] and Feynman and Gell-Mann [34] suggested then in 1957 that the interaction should be of the form \(V-A\). Both groups made a bold proposal contradicting some established experiments. However, with this form they could compute a number of weak processes and the agreement with other experiments was striking.

The \(V-A\) theory is not a renormalizable field theory, nor is it a gauge theory. Feynman and Gell-Mann pointed out that the interaction could be mediated by a heavy vector particle. This proposal was also made by Schwinger, who called the particle \(W\). (Gell-Mann called it \(X\).) Feynman and Gell-Mann concentrated on the hadronic vector current that can be written as (now suppressing the space-time index \(x\))

\[ J^\mu = \overline{p} \gamma^\mu n. \]

By introducing an isospin space where
and Pauli matrices $\tau$ the vector current can be written as

$$J^\mu_h = \overline{N} \gamma^\mu \tau^+ N,$$

and can be compared to the conserved electromagnetic one

$$J^\mu_e = \overline{N} \gamma^\mu \frac{1}{2} (1 + \tau^3) N.$$

They then postulated that the weak vector current is also conserved, the Conserved Vector Current hypothesis, CVC, and pointed out that it should have a common origin with the electromagnetic current. Here it is possible to find thoughts about a unified theory. There should be a symmetry behind the weak current too. Feynman and Gell-mann discuss it in terms of the so-called Sakata Model [35]. Shoichi Sakata had proposed in 1956 that there are three basic hadrons, $p$, $n$ and $\Lambda$, and three basic leptons $e$, $\mu$ and $\nu$. They were then identified with the proton, neutron and the $\Lambda$–particle and the electron, muon and the neutrino. This Nagoya school was quite influential and the Sakata particles finally became the quarks of Gell-Mann, though with a somewhat different purpose. The CVC hypothesis had been suggested earlier by Zel’dovich and Gerstein [36] but had not been gained any attention in the West.

Several scientists now proposed a universal four-fermi interaction for the weak force of the form

$$H_V = G J^*_\lambda J^\lambda,$$

with

$$J^\lambda = J^\lambda_h + J^\lambda_i,$$

where $J^\lambda_h$ is the hadronic current. There were two problems with this description. Firstly, no universal coupling $G$ was found when investigating the hadronic decays. The axial vector term $A$ in $V-A$ coupled with a 26 % stronger coupling, the Gamow-Teller coupling $G_A$, than the vector term $V$. Secondly, there was a discrepancy in the vector coupling when measuring the decay of radioactive oxygen, $^{14}O$. It was not $G$ but $0.97G$. The first problem was addressed in two landmark papers, one by Gell-Mann and Lévy [37] and one by Yoichiro Nambu [38], submitted within four days of each other in 1960. Both suggested that that the axial vector current was only partially conserved, PCAC. Before discussing these papers, we must make a digression.

**Superconductivity**

In 1956, John Bardeen, Leon Cooper and Robert Schrieffer [39] (Nobel Prize 1972) found the long-sought solution to the puzzle of superconductivity. It was based on electrons forming pairs, Cooper pairs, through interaction with the phonons. The pairs being bosons could then condensate and establish a band gap. This theory is based on electromagnetism, so how was the gauge invariance broken and what remained of it? This was the starting point for Nambu
when he wanted to understand the new theory in a field theoretic framework. It took him some two years to understand this problem and to show that the gauge invariance is still there but non-linearly realized. The problem is that the vacuum state is a condensation of Cooper pairs with charge -2 and is hence in principle charged. Nikolay Bogoliubov [41] introduced in 1958 quasi-states that are linear combinations of electrons and holes. They are not eigenstates of the charge operator but can be used to define a new vacuum. Phillip Anderson [42] (Nobel Prize 1977) found furthermore that there are collective excitations in the gap region. Nambu now treated the BCS Hamiltonian in a Feynman-Dyson formulation and showed that by taking into account “radiative corrections” for the vertices, Ward identities could be established (a sign of gauge invariance) and the collective excitations made their appearance as solutions to his vertex equations. He then found that different solutions could be obtained. One of them is the superconducting phase, and Nambu could perform all calculations in a gauge invariant way. For example, the computation of the Meissner effect, the expulsion of magnetic field lines from a superconductor, could be performed in a strictly gauge-invariant way, a computation that had been challenged in the original BCS treatment. Nambu had discovered *spontaneous symmetry breaking* in a field theoretic formulation. The beauty of this breaking is that even though the ground state is non-symmetric, all the virtues of the symmetry are still there in the theory, and calculations can be performed using the restrictions that the symmetry provides.

Spontaneous symmetry breaking had long been known in condensed matter physics as a state of order. Heisenberg [43] had already introduced it in 1928 in his theory for magnetism. Here, the Hamiltonian is rotationally invariant, but the lowest state is a state with the spins ordered in one direction, obviously breaking the rotational symmetry. The new understanding that Nambu provided was that this form of breaking could also be performed in a field theory.

**Spontaneous Symmetry Breaking in Particle Physics**

The really bold assumption that Nambu now made in 1960 [44] was that spontaneous symmetry breaking could also exist in a quantum field theory for elementary particles. In magnetism or in superconductivity, the “vacuum” is really a ground state, in the first case of atoms and in the second case of electrons and atoms. It is possible to give a vacuum expectation value for a physical quantity like a spin. In a particle theory, the vacuum is an abstract state and was assumed to be empty apart from quantum fluctuations. Nambu now introduced vacuum expectation values for certain fields. In fact, he put forward a scheme for the theory of the strong interactions that mimicked superconductivity in the following way:

<table>
<thead>
<tr>
<th>Superconductivity</th>
<th>Strong Interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>free electrons</td>
<td>hypothetical fermions with small mass</td>
</tr>
<tr>
<td>phonon interaction</td>
<td>unknown interaction</td>
</tr>
<tr>
<td>energy gap</td>
<td>observed mass of the nucleon</td>
</tr>
<tr>
<td>collective excitations</td>
<td>mesons, bound states</td>
</tr>
<tr>
<td>charge</td>
<td>chirality</td>
</tr>
<tr>
<td>gauge invariance</td>
<td>chiral invariance, possibly approximate</td>
</tr>
</tbody>
</table>
The chiral invariance corresponds to the axial current and Nambu first used this idea to address the PCAC problem in the paper mentioned above [34]. By also assuming a small explicit breaking of the chiral invariance, he concluded that the pion has a small mass, much smaller than other scales in the problem, and he formulated PCAC and could then compute the so-called Goldberger-Treiman relation [45], which is a relation between the axial vector coupling $G_A$, the decay constant of the pion and the coupling between two nucleons and a pion. Goldberger and Treiman had derived it from dispersion theory with some very brave assumptions. Here, it followed naturally from Nambu’s ideas. He followed up on his ideas in two papers with Giovanni Jona-Lasinio [46] and with other collaborators [47]. Since there was no theory for the strong interactions, they discuss two possible cases. In the first, they use a theory with only fermions and in the other, they discuss a theory with intermediate vector particles. Nambu was well aware of Feynman’s functional formulation of quantum field theories and had in fact worked on it in Japan around 1950. He realized that the fermion model is only an effective theory with other physical fields integrated out, but that it will give similar results to a full theory like the second one. He came back to a complete theory a few years later as described below. Nambu and collaborators made model independent calculation and could test their ideas directly. In the case of “soft” pions (pions with small momentum transfer) they could compute cross sections for scattering between pions and nucleons that agreed with the experimental data.

Note that Nambu’s picture was very different from Yukawa’s. In the latter, the pion was the intermediary between the nucleons. In Nambu’s picture, the pion is a bound state of constituent particles with small masses. This picture is four years before Gell-Mann’s [48] and George Zweig’s [49] introduction of quarks. The pion is also the sign of the chiral symmetry breaking.

Nambu’s treatment of the BCS theory is non-relativistic since there is a Fermi surface. The particle physics case is treated relativistically. The pion is a composite particle and it is this composite that becomes massless when there is no explicit breaking of the symmetry. In 1961, Jeffrey Goldstone [50] performed a similar calculation with a scalar field, getting a vacuum expectation value, and showed that this also leads to a massless particle in the spectrum, a Nambu-Goldstone boson. In his treatment of the BCS theory, Nambu furthermore found that there is a state with energy and momentum zero, a “massless” phonon. When he takes the Coulomb field into account, these are transformed into “massive” plasmons. Note that the chiral symmetry above is a global symmetry and when a global symmetry is spontaneously broken, a massless particle appears. In superconductivity, the symmetry is a gauge symmetry and this leads to massive states. The same ideas were tried in 1964 for relativistic gauge theory by Robert Brout and François Englert [51] and also by Peter Higgs [52]. They found that a spontaneously broken gauge symmetry, as in the non-relativistic version of Nambu, does not produce a massless particle. Instead, this mechanism gives the vector field a mass and a scalar state, the still today hypothetical Higgs particle, which is also a characteristic feature of such a theory.

In 1965, Nambu, together with M.-Y. Han [53], suggested that the underlying theory for the strong interactions should be a non-abelian gauge theory based on the gauge group $SU(3)$. This was a year after Gell-Mann’s work on quarks and Han and Nambu chose to have quarks with integer electric charges. They now had also a new quantum number that was to be called colour. Their choice of electrical charges turned out to be incorrect, but apart from that the idea was essentially correct. In 1971, Gerhard ’t Hooft, together with Martinus Veltman [54] (Nobel Prize 1999), proved that non-abelian gauge theories are renormalizable even if the gauge symmetry is spontaneously broken. This started the modern era of the Standard Model
of particle physics. It then became clear that a non-abelian gauge theory with the gauge symmetry $SU(2) \otimes U(1)$ for the weak and electromagnetic interactions proposed by Steven Weinberg in 1967 [55] and Abdus Salam [56] in 1968, based on earlier work by Sheldon Glashow [57] could indeed be a viable model for Nature (Nobel Prize to the three, 1979). The scales in the model come via spontaneous symmetry breaking. In fact the electro-weak theory of Glashow, Salam and Weinberg has now been tested to very great accuracy at the LEP accelerator and all the data supports a spontaneously broken gauge theory [58]. Nambu’s idea of spontaneous symmetry is one of the pillars of the model.

In 1973, it was shown by David Gross and Frank Wilczek [59] and David Politzer [60] that a non-abelian gauge theory shows asymptotic freedom, i.e. the effective coupling constant goes to zero at large energies, contrary to the behaviour in an abelian gauge theory like QED (Nobel Prize to the three 2004). Gross and Wilczek [61] then proposed a non-abelian gauge theory like the one proposed by Han and Nambu, but with fractionally charged quarks, as the theory of the strong interactions. This came to be called Quantum ChromoDynamics (QCD). The possibility of having quarks in a triplet representation interacting with a vector particle had been discussed the year before by Harald Fritzsch and Gell-Mann [62] in an attempt to catalogue all possible models. The Standard Model for Particle Physics was born based on the gauge symmetry $SU(3) \otimes SU(2) \otimes U(1)$, with the symmetry spontaneously broken to $SU(3) \otimes U(1)$. This model has in the last thirty years been verified with exquisite precision, and it seems clear that the Standard Model is the correct model for the energy scale at which we can measure today.

Even though QCD is the correct theory for the strong interactions, it cannot be used to compute at all energy and momentum scales. For many purposes, the originally idea of Nambu and Jona-Lasinio works better. Chiral Perturbation Theory is essentially their programme and is an effective theory constructed with a Hamiltonian consistent with the (approximate) breaking of the chiral symmetry of QCD [63]. It allows us to study the low-energy dynamics of QCD, a region where perturbative methods do not work for QCD. Since it is an effective theory in the way Nambu and Jona-Lasinio originally set up their programme, the degrees of freedom are not the quarks and the gluons but the experimentally detected hadrons. This model has been the standard tool to compute strong interaction processes at this energy range and has been of the utmost importance in recent years for interpreting data, both for processes involving mesons and for heavy ion collisions.

Spontaneous symmetry breaking plays an important role in many fields of science. A phase transition is often due to spontaneous breaking of a symmetry. This is a common phenomenon in condensed matter physics, in cosmology and also in chemistry and biology. The Nambu-Goldstone boson in this case is translated into a relation between the energy and the momentum, which says that for small momenta the energy tends to zero.

Quark Mixing

As mentioned above, there was also a problem with the universality of the vector coupling found in the decay of radioactive oxygen, $^{14}O$. A value of $0.97\, G$ had been measured for that process. Gell-Mann and Lévy [37] discuss various explanations for it but put eventually forward the idea that it could be that the contributions of $\Delta S = 0$- and $\Delta S = 1$ to the hadronic current, where $S$ is the strangeness, have different strengths. They discuss this in terms of the Sakata Model and write for the vector part
and likewise for the axial part. This amounts to a unitary transformation of the \( n \) and \( A \)-states.

With a value \( \varepsilon^2 = 0.06 \), they find a correct value for the coupling strength of the decay. This was an ingenious proposal based on only one experiment. Others had previously suggested that renormalization effects could be at work, but Gell-Mann and Lévy wanted to keep the coupling universal and find other ways out. They believed that a renormalizable theory must be behind, with a universal coupling constant.

The Sakata Model had been important and as long as currents bilinear in fields were discussed, it was useful. The model did treat mesons and baryons differently. In another ground-breaking step in 1961, Gell-Mann [64] kept the results for the currents and postulated that the particles should transform as representations of a (broken) \( SU(3) \). By using the lowest representation \( 3 \), \( 3 \otimes 3 = \mathbf{8} \oplus \mathbf{1} \) is obtained. Gell-Mann now discards the Sakata particles and postulates that all mesons belong to the representation \( \mathbf{8} \), “the eight-fold way”. Similarly, the lowest lying baryons belong to an \( \mathbf{8} \). Yuval Ne’eman [65] also puts forward this proposal. All known hadrons were found to belong to representations of this group. The symmetry was broken, but in a controlled way, and Gell-Mann used it to propose missing particles that were subsequently found. He now left the idea of fundamental particles, at least for the time being. In an article [66] the year after, he continued this line and argued that the \( SU(3) \) symmetry is universal. The weak and the electromagnetic currents must transform as components of an octet. There is one octet of vector currents and one of axial vector currents. From these assumptions, he could derive a series of selection rules such as the absence of processes with \( \Delta S = -\Delta Q \), where \( S \) is strangeness and \( Q \) is the charge.

In 1959, the large particle accelerators at Brookhaven and CERN came into operation and a great deal of new experimental data was obtained. A much clearer picture of the weak decays of the hadrons was reached, and with these results, Nicola Cabibbo [67] in 1963 made a very important contribution. He took as his starting points three assumptions from Gell-Mann’s earlier work:

- The weak current \( J_h \) transforms as a component of an octet under \( SU(3) \).
- The vector part is part of the same octet as the electromagnetic current.
- The weak current is universal and of length 1.

The first two had been the essence of Gell-Mann’s paper from the year before and the third is the same as the proposal by Gell-Mann and Lévy. He introduces an angle instead of their parameter \( \varepsilon \) through

\[
\cos \theta = \frac{1}{(1 + \varepsilon^2)^{1/2}},
\]

\[
\sin \theta = \frac{\varepsilon}{(1 + \varepsilon^2)^{1/2}}.
\]

He can then use a general form for the hadronic current using \( SU(3) \) language.

\[
J_\mu = \cos \theta [V^{\mu 1} + iV^{\mu 2} - (A^{\mu 1} + iA^{\mu 2})] + \sin \theta [V^{\mu 4} + iV^{\mu 5} - (A^{\mu 4} + iA^{\mu 5})].
\]

This is written in Gell-Mann’s \( SU(3) \) language. The matrices \( V \) and \( A \) are Gell-Mann’s \( \lambda \)-matrices, which are eight \( 3 \times 3 \) matrices representing the group. They are a generalization of
Pauli’s $\sigma$-matrices. He used this current to compare the branching ratios of the kaon and the pion and found a value in agreement with the one found by Gell-Mann and Lévy. He then continued to compute branching ratios for strange baryons and found that the value of the angle $\theta$ seemed to be universal. The Cabibbo Theory with the Cabibbo angle $\theta$ quickly became a standard framework for the weak interactions. It turned out to be universal and an ever-increasing multitude of data could be fitted into it. It has been a cornerstone of weak interactions.

In 1964, Gell-Mann [48] came back to the idea of fundamental constituents from which one could construct both the hadrons and the hadronic currents. He introduced three quarks, the u-quark, the d-quark and the s-quark with the electric charges $2/3$, $-1/3$ and $-1/3$. These transform as a triplet under $SU(3)$. Their antiparticles transform as $\bar{3}$. The mesons are built up by a quark-antiquark pair and the hadrons by three quarks. This idea was also put forward by George Zweig [49]. Gell-Mann used a free quark model to extract the electromagnetic as well as the vector and axial vector currents in terms of the quark field and they then got the natural form

\[ j_\mu^e = \frac{2}{3} u \gamma^\mu u - \frac{1}{3} d \gamma^\mu d - \frac{1}{3} s \gamma^\mu s, \]

\[ j_\mu^b = \cos \theta \bar{u} \gamma^\mu (1 - \gamma_5) d + \sin \theta \bar{u} \gamma^\mu (1 - \gamma_5) s. \]

The hadronic current now looks again like that of Gell-Mann and Lévy. He used these currents to extract the commutation relations for the physical currents and this led to many results that led up to the Standard Model. Once it was discovered, these relations could be put on a firm ground.

**Neutral Kaon and CP-Invariance**

After having introduced the quantum number strangeness, Gell-Mann, together with Abraham Pais [68] in 1955, pointed out that kaons must have unusual properties. (The discussion was in terms of $\theta$- and $\tau$-particles, but we use modern language here.) The kaons came as two isodoublets ($K^+$, $K^0$) and their antiparticles ($K^-$, $\bar{K}^0$) with strangeness $+1$ and $-1$. Under strong interactions, the two neutral particles are different, since this force respects strangeness, but under the weak force they can be mixed since this force does not respect strangeness. Gell-Mann and Pais found that it was more advantageous to use a mixture of these particles. Since it was believed that the charge conjugation was respected by all forces, eigenstates of the $C$ operator should be used. They introduced the states

\[ K^0_+ = \frac{1}{\sqrt{2}} (K^0 + \bar{K}^0), \]

\[ K^0_- = \frac{1}{\sqrt{2}} (K^0 - \bar{K}^0). \]

These are eigenstates with the eigenvalues $+1$ and $-1$ if the arbitrary phase in the $C$-operation is chosen as $\bar{K}^0 = CK^0$.

These two neutral states should then have different decay modes and hence different lifetimes. In the decay $K^0 \rightarrow 2\pi$, the pions must have a relative angular momentum which is the same as the spin of the kaon, since the pions do not have any spin. A neutral pair of two pions is an
eigenstate of $C$. Hence only one of the $K^0_\pm$ can decay into two pions, namely $K^0_+$. The other neutral kaon $K^0_0$ should then have a longer lifetime since the phase space to decay into three pions is smaller than the one for decay into two. This led to a series of detectable phenomena [69]. A new naming of $K^0_0$ and $K^0_0$ as $K^0_S$ and $K^0_L$ was now accepted. This very successful picture was questioned after Lee’s and Yang’s suggestion that the weak force does not conserve parity. They also suggested that the $C$ invariance is broken in the weak decays. Landau [32], however, rescued the picture by pointing out that $CP$ should be a good symmetry and that by exchanging the $C$ operation with a $CP$ operation, the results should still hold.

Non-Conservation of $CP$

The picture above worked well for another seven years. Then in 1964, James Cronin, Val Fitch and collaborators [70] (Nobel Prize 1980 to Cronin and Fitch) announced the very surprising result that they had found the process $K^0_L \to \pi^+ \pi^-$. The decay mode was found to be small fraction $2 \times 10^{-3}$ of all charged decay modes. This was totally unexpected. All theories that had been constructed were invariant under the $CP$-symmetry. Since there was very strong evidence that the combined $CPT$ symmetry must be conserved, the breaking of $CP$ indicated that the $T$ invariance also had to be broken for this process. For quite some time, $CP$ breaking was only found in kaon physics. To keep the formalism of Gell-Mann and Pais, it was necessary to give the coefficients in their mixing formula small imaginary contributions.

$$K^0_S = \frac{1}{\sqrt{2(1 + \varepsilon^2)}}((1 + i\varepsilon)K^0 + (1 - i\varepsilon)\overline{K^0}),$$
$$K^0_L = \frac{1}{\sqrt{2(1 + \varepsilon^2)}}((1 + i\varepsilon)K^0 + (1 - i\varepsilon)\overline{K^0}).$$

There were quickly several suggestions how to explain this new effect. Lincoln Wolfenstein [71] introduced a superweak $\Delta S = 2$ interaction which could explain the effect. It only affected decays of kaons, so any other decay which violates the $CP$ invariance would refute this proposal. It would take some 35 years before this was detected in $B$-decays.

$CP$-Transformations in a Quantum Field Theory

We see from the formula above that to get $CP$-violating effects in a quantum field theory we must have complex parameters. We do have complex fields but also need complex coupling constants. A $CP$ transformation of a complex field can be written as

$$U_{CP} \phi U_{CP}^{-1} = e^{i\alpha} \phi^*,$$

where $\alpha$ is an arbitrary angle. Suppose now that the interaction term is of the form

$$H_v = g\phi^*O + g^*\phi^*O^*,$$

where $g$ is a coupling constant and $O$ an operator bilinear in the field. Let us now perform a $CP$ transformation on this term
By choosing the angle such that

$$\alpha = -2 \arg(g),$$

it is found that the interaction is invariant even though we started with a complex coupling constant. Hence, in order to have a CP violation more fields and more couplings are needed. We must check if the couplings stay complex when we perform field redefinitions and use the freedom in the CP transformations. From this we see that we cannot have CP violation in the Gell-Mann-Lévy-Cabibbo theory. There must be more fields.

The Work of Kobayashi and Maskawa

After the tumultuous breakthrough in 1971, the particle physics world became focused on the Glashow-Salam-Weinberg Model. At last, there was a quantum field theory that satisfied all the requirements that the previous research had set up for such a theory. It was also seemingly consistent, and quantum corrections could be computed. It was, of course, not clear that this was the correct model. Within a year or so, many possible such theories were constructed that all agreed with the experimental data of the day. They only differed in the outcomes of higher quantum corrections which so far it had not been possible to test. It was at this time in 1972 that two young Japanese physicists from Sakata’s Nagoya School, Makoto Kobayashi and Toshihide Maskawa [72] addressed the problem of CP violation, this time in terms of the Glashow-Salam-Weinberg Model. Since they knew of the analysis above, they realized that they had to extend the model to get a term with a true complex coupling constant. They concentrated on the coupling of the W-bosons to the quark fields. It is schematically of the form

$$gW^\nu \bar{q}_i \gamma_\nu (1 - \gamma_5) q_j.$$

Can such a term have a complex coupling constant? They started by investigating models with two families of quarks. They found that there was no way to get a complex coupling constant however they tried to play with different representations under the SU(2) symmetry group, unless they added new scalar fields. Finally, they investigated a model with three families with SU(2) quantum numbers, generalizing the ones that Salam and Weinberg used for the quarks u, d and s

$$\begin{pmatrix} u \\ d \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}.$$

They then found that via a unitary transformation of the quarks d, s, and q_3 a coupling with a free phase is obtained that cannot be absorbed in another term.

This is in fact a result known in mathematics since around 1950, but the contacts between mathematics and physics were not great around 1970. Suppose there are two vectors with N components
\[ Q = \begin{pmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_n \end{pmatrix}, \quad q = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix}, \]

connected via a unitary matrix \( U \) as \( \bar{Q} U q \). It is then possible to make a so-called Iwasawa decomposition [73] such that \( U = D_1 V_{CKM} D_2 \). The matrices \( D_1 \) and \( D_2 \) are diagonal and each contain \( n \) phases. These can be absorbed in the quark fields. One of them is an overall phase that is not observable. The counting of degrees of freedom is then: \( V_{CKM} \) and \( U \) have \( n^2 \) degree of freedom. The diagonal matrices \( D \) take away \( 2n-1 \) degrees of freedom. Hence the remaining degree of freedom for \( V_{CKM} \) is \( n^2 - (2n-1) = (n-1)^2 \). When \( n=2 \), the result is 1 as in the case of the Gell-Mann-Lévy-Cabibbo theory. In the case \( n=3 \), there are four degrees of freedom which can be expressed as three angles and a phase. In a three-dimensional space, there are only three independent angles. Hence the fourth degree of freedom has to be imaginary. Kobayashi and Maskawa wrote the remaining matrix as

\[
V_{CKM} = \begin{pmatrix}
c_1 & -s_1 c_3 & -s_1 s_3 \\
s_1 c_2 & c_1 c_2 c_3 + s_2 s_3 e^{i\delta} & c_1 c_2 s_3 + s_2 c_3 e^{i\delta} \\
s_1 s_2 & c_1 s_2 c_3 + c_2 s_3 e^{i\delta} & c_1 s_2 s_3 - c_2 c_3 e^{i\delta}
\end{pmatrix},
\]

where \( s_1 = \sin \theta_1, c_1 = \cos \theta_1 \), etc. and the explicit phase \( \delta \) is seen. The angle \( \theta_i \) is the Cabibbo angle. When discussing quark mixing, the matrix is usually called the Cabibbo-Kobayashi-Maskawa (CKM) matrix. Nowadays, the matrix is written in a slightly different form, recommended by the Particle data Group [58],

\[
V_{PDG} = \begin{pmatrix}
c_{12} c_{13} & s_{12} c_{13} & s_{13} e^{-i\delta} \\
-s_{12} c_{23} - c_{12} s_{23} s_{13} e^{i\delta} & c_{12} c_{23} - s_{12} s_{23} s_{13} e^{i\delta} & s_{23} c_{13} \\
s_{12} s_{23} - c_{12} c_{23} s_{13} e^{i\delta} & -c_{12} s_{23} - s_{12} c_{23} s_{13} e^{i\delta} & c_{23} c_{13}
\end{pmatrix},
\]

where \( s_{ij} = \sin \theta_j, c_{ij} = \cos \theta_j, i, j = 1,2,3 \). Again, the fact that the three-family theory contains \( CP \) violation is witnessed by the presence of a phase angle that cannot be transformed away.

A convenient expression summarizing the condition for \( CP \) violation using a commutator of products of mass matrices was given in 1985 by Cecilia Jarlskog [74].

### The Experimental Verification of the Kobayashi-Maskawa Theory

The Kobayashi-Maskawa paper was submitted on Sept 1, 1972. At that time, only three quarks were known. Their theory contained six quarks and was not particularly noticed at the time. There was strong evidence for a fourth quark which had been put forward in 1970 by Sheldon Glashow, Jean Iliopoulos and Luciano Maiani [75]. There was a problem with the then existing theory for the weak interactions. It led to strangeness changing neutral currents which were not seen in experiments. The three authors suggested the introduction of a fourth quark, which came to be called the charm quark, \( c \). By introducing this quark, two full
families are obtained and strangeness changing neutral currents are indeed suppressed. In 1974, a new, quite heavy particle, the $J/\psi$-particle, was indeed found by Samuel Ting et al [76] and Burton Richter et al. [77] (Nobel Prize to Richter and Ting 1976). It was rather quickly understood as a $\bar{c}c$ state. Particles with a charm quantum number were discovered a few years later. Signs of a new heavy lepton started to come in 1975 and the discovery was established in 1977 [78] (Nobel Prize to Martin Perl 1995). This indicated a third family of leptons. At this stage, the Kobayashi-Maskawa paper started to come into focus, at the beginning for model building but soon also for phenomenological purposes. Now the evidence that the Glashow-Salam-Weinberg Model was indeed the correct one for the weak and electromagnetic interaction was also accumulating. Also in 1977, Leon Lederman (Nobel Prize 1988) and his group found the fifth quark, the $b$-quark [79]. It was not until 1994 that the sixth quark, the $t$-quark, was discovered [80].

The discovery of the $b$-quark and its long lifetime led to new possibilities to test the $CP$ violation and to choose between the KM Model and Wolfenstein’s proposal. The $b$-quark has the same quantum numbers as the $s$-quark apart from its inherent quantum number. Neutral $B$-mesons (with one $b$-quark) must then have the same $CP$ properties as the kaons. In fact, it was argued that the violation should be quite large for non-leptonic decays in the KM Model. This led to the setting up of “$B$ factories” at SLAC at Stanford and KEK in Japan. The respective collaborations BABAR and BELLE have now measured the $CP$ violation in remarkable agreement with the model [81] and all experimental data are now in impressive agreement with the model [58]. The model of Wolfenstein was also ruled out with strong significance by experiments measuring direct $CP$ violation, at Fermilab and CERN in 2002–3 [82]. It has been shown that Nature follows the Kobayashi-Maskawa Model to describe the weak interactions in particle physics. The fundamental constituents of Nature come in three families, at least at the energies we can measure, and that allows for a $CP$ violation distinguishing matter and antimatter. It should be mentioned too that the model also passes all theoretical checks that physicists have set up.

In 1967, Andrei Sakharov [83] (the Nobel Peace Prize 1975) pointed out in a famous work that $CP$ violation must be the cause of the asymmetry in the universe. It contains more matter than antimatter. The $CP$ violation that the KM Model gives rise to is most probably not enough to explain this phenomenon. To find the origin of this $CP$ violation we probably have to go beyond the Standard Model. Such an extension should exist for other reasons as well. It is believed that at higher energies other sectors of particles, so heavy that the present day accelerators have been unable to create them, will augment the model. It is natural that these particles will also cause $CP$ violations and in the tumultuous universe just after the Big Bang these particles could have been created. These particles would have been part of the hot early universe and could have influenced it, by an as yet unknown mechanism, to be dominated by matter. Only future research will tell us if this picture is correct.
References


