# 129A Lecture Notes
## Weak Interactions I

## 1 Nuclear $\beta$-decay

The nuclear $\beta$-decay caused a great deal of anxiety among physicists. Both
$\alpha$- and $\gamma$-rays are emitted with discrete spectra, simply because of the energy
conservation. The energy of the emitted particle is the same as the difference
between the initial and final state of the nucleus. It was much more difficult
to see what was going on with the $\beta$-decay, the emission of electrons from
nuclei.

Chadwick once reported that the energy spectrum of $\beta$ electrons is con-
tinuous. The energy could take any value between 0 and a certain maximum
value. This observation was so bizzarre that many more experiments fol-
lowed up. In fact, Otto Han and Lise Meitner, credited for their discovery
of nuclear fission, studied the $\beta$ spectrum and claimed that it was discrete.
They argued that the spectrum may appear continuous because the electrons
can easily lose energy by breamsstrahlung in material. The maximum energy
observed is the correct discrete spectrum, and we see lower energies because
of the energy loss. The controversy went on over a decade. In the end a
definitive experiment was done by Ellis and Wooseley using a very simple
idea. Put the $\beta$-emitter in a calorimeter. This way, you can measure the
*total* energy deposit. They have demonstrated that the total energy was
about a half of the maximum energy on average. The spectrum is indeed
continuous.

The fact that the $\beta$-spectrum is continuous was su puzzling to people,
and let Niels Bohr to say even this:

> At the present stage of atomic theory, however, we may say that
> we have no argument, either empirical or theoretical, for uphold-
> ing the energy principle in the case of $\beta$-ray disintegrations.

He was ready to give up the energy conservation! This quote shows how
desperate people were.

Pauli, also desperate about this problem, wrote a letter to people who
attended a meeting in Tübingen in December 1930. He himself could not go
to the meeting because he had to attend a ball. Nonetheless, he could not

resist to propose a speculation to solve this problem even though somewhat reluctantly. Here is a quote from his letter:

```
4th December 1930

Dear Radioactive Ladies and Gentlemen,

As the bearer of these lines, to whom I graciously ask you to listen,
will explain to you in more detail, how because of the "wrong"
statistics of the N and Li6 nuclei and the continuous beta spectrum, I
have hit upon a desperate remedy to save the "exchange theorem" of
statistics and the law of conservation of energy. Namely, the
possibility that there could exist in the nuclei electrically neutral
particles, that I wish to call neutrons, which have spin 1/2 and obey
the exclusion principle and which further differ from light quanta in
that they do not travel with the velocity of light. The mass of the
neutrons should be of the same order of magnitude as the electron mass
and in any event not larger than 0.01 proton masses. The continuous
beta spectrum would then become understandable by the assumption that
in beta decay a neutron is emitted in addition to the electron such
that the sum of the energies of the neutron and the electron is
constant...
```

Remember this letter was written before the discovery of the neutron by Chadwick in 1932. The picture of $^{14}$N was that it composed of 14 protons and 7 electrons. On the other hand, the molecular spectrum showed that $^{14}$N was a boson. He attempted to hit two birds with one stone: this statistics puzzle and the apparent energy non-conservation. He proposed that the $^{14}$N is made of 14 protons, 7 electrons, and 7 neutrinos. The $\beta$-decay emits one electron and one neutrino at the same time. Because you don't detect the neutrino, by assumption, it appears that the energy is not conserved. On the other hand, this composition makes the $^{14}$N a boson.

When Chadwick discovered the neutron in 1932, the picture of the nuclei was established fairly quickly. $^{14}$N is made of 7 protons and 7 neutrons, and there is no statistics puzzle. But what about the $\beta$-decay? It was clear that Chadwick's neutron is not Pauli's "neutron" which was supposed to be at least as light as the electron. It was Fermi who formalized the Pauli's idea into the first theory of weak interaction. First of all, he renamed

Pauli's "neutron" *neutrino*, the little neutral guy, being Italian. He further postulated that a neutron can disintegrate into a proton, an electron, and an electron-anti-neutrino.

The interaction he postulated is modeled after the electromagnetism. The Feynman rule of the Quantum ElectroDynamics is that the photon couples to the electric current. In the non-relativistic notation, the scalar potential couples to the charge density, $\psi^\dagger \psi$, while the vector potential to the current density, $\psi^\dagger \dot{\vec{x}} \psi$. Using the Dirac equation, we know that $\dot{\vec{x}} = \frac{1}{i\hbar}[\vec{x}, H] = \frac{1}{i\hbar}[\vec{x}, \vec{\alpha} \cdot \vec{p}] = \vec{\alpha}$. Therefore, the four-vector current is $(\psi^\dagger \psi, \psi^\dagger \vec{\alpha} \psi)$. Using the relativistic notation $\gamma^0 = \beta$, $\gamma^i = \beta \alpha^i$, and $\bar{\psi} \equiv \psi^\dagger \gamma^0$, it is rewritten as $\bar{\psi} \gamma^\mu \psi$. The electron-photon vertex for the interaction is given by $-ie\gamma^\mu$. Similarly, Fermi imagined that the same is true with the weak interaction, and wrote the amplitude for the neutron decay,

$$G_F \bar{u}_p \gamma^\mu u_n \bar{u}_e \gamma_\mu v_\nu. \tag{1}$$

Here the neutrino wave function is the negative energy solution because the final state is the anti-neutrino. He has found that $G_F \sim 10^{-5} \text{GeV}^{-2}$, a very small number, justifying the name the "weak" interaction. It actually suggests that the weak interaction is weak because it is short-ranged, acting only over the distance $\hbar c \sqrt{10^{-5} \text{GeV}^{-2}} \sim 6 \times 10^{-17}$ cm.

Back to the non-relativistic limit, the dominant piece in this interaction is the time component, becaue the spatial components are suppressed by the velocity of nucleons. Because the time component is just $\bar{u}_p \gamma^0 u_n = u_p^\dagger u_n$, it just replaces the neutron in a nucleus by a proton. Therefore this transition does not change the spin of the nucleus, nor the parity. This type of transition is called Fermi transition. For example, the inverse beta decay $^{14}\text{O} \rightarrow {}^{14}\text{N}^* e^+ \nu_e$ is a $0^+$ to $0^+$ transition, where $^{14}\text{N}^*$ is an excited state that belongs to the same isospin one multiplet as the ground states of $^{14}\text{O}$ and $^{14}\text{C}$. Even though the phase space is much larger for the decay to the ground state of $^{14}\text{N}$, the $\beta$-decay is almost 100% into this excited state. This suggests an important point: the $\beta$-decay is caused by the raising or lowering operator in the isospin space.

Another important point is that $^{14}\text{C}$ cannot decay into the isospin partner $^{14}\text{N}^*$ because the latter has higher energy. It decays into the $1^+$ ground state of $^{14}\text{N}$ with a long lifetime. That is why $^{14}\text{C}$ is useful for carbon dating. The point is that $0^+ \rightarrow 1^+$ transition is possible. Such transitions are called Gamov–Teller transitions. For this to be possible, the original

Fermi's interaction is not complete. It has to be supplemented by something else. Considering all possible 16 Dirac matrices, $\bar{u}_p u_n$ (S) and $\bar{u}_p \gamma^\mu u_n$ (V) reduce to the Fermi transition in the non-relativistic limit, while $\bar{u}_p \sigma^{\mu\nu} u_n$ (T) and $\bar{u}_p \gamma^\mu \gamma_5 u_n$ (A) to the Gamov–Teller transition $u_p^\dagger \vec{\sigma} u_n$. The last possibility $\bar{u}_p \gamma_5 u_n$ (P) does not cause any transitions in the non-relativistic limit. Therefore, it is clear that you need either S or V, and T or A. The true answer turned out to be $V$ and $A$ as we will see below.

# 2  Parity Violation

As strange the strange particles were, one very puzzling feature emerged. One particle, called $\tau^+$, decays into three pions $\pi^+\pi^+\pi^-$ or $\pi^+\pi^0\pi^0$. Another one, called $\theta^+$, decays into two pions $\pi^+\pi^0$. Both are spin zero particles of strangeness one. The analysis of the final state showed that the $\tau^+$ decays into a parity odd state, while the $\theta^+$ into a parity even state. The mystery was that both particles appeared to have the same mass and the same lifetime. Could this just be a coincidence?

   T.D. Lee and C.N. Yang pointed out in 1956 that maybe these two particles could be the same particle.[1] Of course it is possible only if the parity is not preserved in these decays. They examined carefully the available evidence for parity conservation, and concluded that there were many good evidence for parity conservation in the strong and the electromagnetic interactions, while there was none in the weak interaction. They further proposed various ways the parity (non)conservation can be tested experimentally in the weak interaction.

   C.S. Wu did such an experimental test quickly. Read Chapter 6 of Cahn–Goldhaber about the experiment. She has observed a correlation between the spin of $^{60}$Co nucleus and the emitted $\beta$-electrons. If the parity were conserved, a final state and its parity conjugate must have the same probability. Under parity, the spin does not change, while the direction of the electron flips. Therefore there should not be any correlation between the direction of the spin and the direction of the electron. She observed that there was a clear correlation in a beautiful experiment. (This paper is a must-read!) The parity, which was long believed to be the true symmetry of nature, fell in 1957.

---

[1]See anecdote that involves Martin Bloch, Richard Feynman and Lee–Yang in Cahn–Goldhaber, Chapter 6.

The words spread quickly that the parity is violated, and the violation is *large*. Hearing this, R. L. Garwin, L. M. Lederman, and M. Weinrich, decided to test parity violation in muon decay. According to Lederman's book "The God Particle," his student was mounting his thesis experiment at a cyclotron. Garwin, Lederman, and Weinrich, all excited, stromed the lab, dismounted poor student's experiment, and set up the muon decay experiment quickly. There, they found a strong correlation between the muon spin and the electron direction. Their papers are publisehd next to each other in Physical Review.

Around the same time, there was a big confusion if the nuclear beta decay was due to the S, T combination or V, A combination. It was once "established" that it was due to the S, T combination, that turned out to be false. In the end it was decided that V and A are the correct form of the weak interaction.

What combination of V and A, then? An extremely clever experiment was performed by M. Goldhaber, L. Grodzins, and A. W. Sunyar, to determine the helicity of the neutrino emitted in the nuclear beta decay. It sounds like an impossible task because you don't see neutrinos. How can you measure the helicity of the particle you don't see? They made a very clever choice of the parent and daughter nuclei so that the daughter nucleus undergoes $\gamma$-decay and the helicity of the photon basically tells you what the helicity of the neutrino is. They have found the result that was consistent with 100% left-handed neutrino. Because the helicity and the chirality are in one-to-one correspondence in the ultra-relativistic limit, it follows that the left-handed helicity of neutrinos means the chirality $\gamma_5 = -1$. In other words, the projection operator $(1 - \gamma_5)/2$ is need. (This experiment used the capture of an orbital electron by the nucleus, and hence the net reaction is $e^- p \rightarrow \nu_e n$.) This point determines that the combination of V ($\bar{u}_\nu \gamma^\mu u_e$) and A ($\bar{u}_\nu \gamma^\mu \gamma_5 u_e$) to be of the $V - A$ form $\bar{u}_\nu \gamma^\mu (1 - \gamma_5) u_e$. Because both $V$ and $A$ appear with 50:50 mixture, this form of the interaction is said to be "maximal violation of parity." If $V$ dominates over $A$, you can treat $A$ as a small perturbation that violates parity, and vice versa. But they come with 50:50 mixture, and neither of them is "small," hence the maximal violation. Parity can't be violated more!

Another important point is this. It is only the left-handed chirality particles that are involved in the weak interaction. This point is very important in the later construction of the $SU(2) \times U(1)$ gauge theory.

If neutrinos are indeed all left-handed, it implies that they are exactly

massles. The argument goes by the contradiction. Suppose they are massive, but yet all left-handed. Because they are massive, their speed is always less than the speed of light. Then you can go faster than them, and look back. The neutrinos are moving backward, but still spinning the same way. Then they are right-handed, and hence contradiction. 100% left-handed neutrinos in $\beta$-decay therefore implies exactly massless neutrinos.

The whole argument leads to the following picture. The right-handed neutrinos don't exist. There are both left-handed and right-handed (in the sense of chirality, not the helicity) electrons exist, but only the left-handed electrons participate in the weak interaction. In general, *only left-handed particles and right-handed anti-particles take part in the weak interaction*.

The muon decay $\mu^- \to e^- \bar{\nu}_e \nu_\mu$ then has the amplitude

$$\frac{G_F}{\sqrt{2}} \bar{u}_e \gamma^\mu (1 - \gamma_5) v_{\nu_e} \bar{u}_{\nu_\mu} \gamma_\mu (1 - \gamma_5) u_\mu. \tag{2}$$

(We need a factor of $\sqrt{2}$ to reconcile the original Fermi's purely vector interaction and the final $V - A$ interaction.) This form of the muon decay amplitude is tested very well.

Another immediate test of the $V - A$ theory is the pion decay. The charged pion decays $\pi^- \to \mu^- \bar{\nu}_\mu$ with 99.98770% probability, while the decay $\pi^- \to e^- \bar{\nu}_e$ is possible only at $1.230 \times 10^{-4}$ level. This is counter intuitive. The decay into the electron has much bigger phase space, as the masses of the charged pion (140 MeV) and the muon (105 MeV) are not very different. Why? The $V - A$ nature of the weak interaction explains this point beautifully. Imagine the limit where the electron mass is zero. This is a good approximation as it is much lighter than the pion. If the electron were exactly massless, the chirality and the helicity match, and only the left-handed electron can be produced in the weak decay of the pion. Of course the anti-neutrino must be purely right-handed. Then their spins add up to one. However, the pion does not have any spin, and the spin one final state is not allowed by the angular momentum conservation. Therefore, the decay $\pi^- \to e^- \bar{\nu}_e$ is impossible in the massless electron limit. The decay amplitude is actually proportional to $m_e$, a very small number. On the other hand, the muon is heavy enough in the pion decay and the mismatch between the chirality and helicity allows the muon mode of the decay despite a smaller phase space. The ratio of two
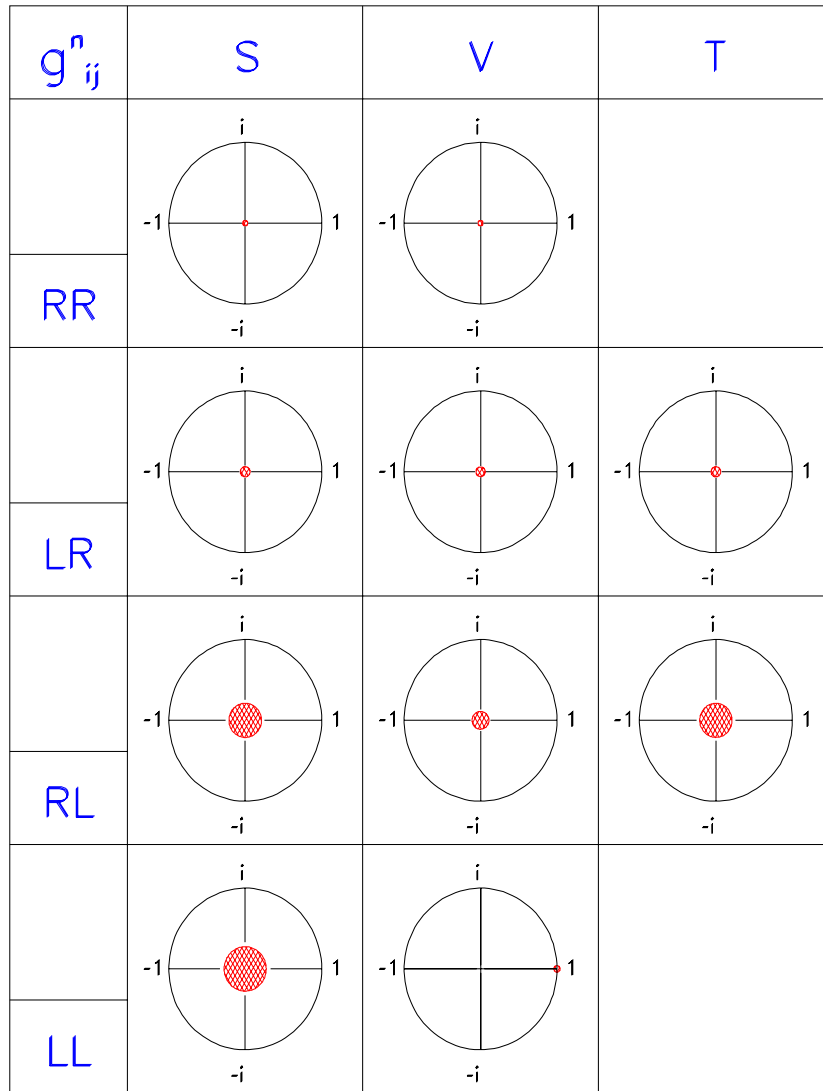
Figure 1: Test of $V - A$ nature in the muon decay. 90% CL experimental limit on possible interactions. Taken from A. Pich, `http://www.arxiv.org/abs/hep-ph/9701263`.

decay probabilities is well-understood as

$$\frac{BR(\pi^- \to e^- \bar{\nu}_e)}{BR(\pi^- \to \mu^- \bar{\nu}_\mu)} = \left(\frac{m_e}{m_\mu}\right)^2 \frac{(1 - m_e^2/m_\pi^2)^2}{(1 - m_\mu^2/m_\pi^2)^2}. \tag{3}$$

This observation demonstrates both the $V - A$ nature of the interaction and also the universal strength between the electron and muon coupling.

In fact, the universal strength of the weak interaction became more and more apparent. The nuclear $\beta$-decay, pion decay (both into electron and muon), muon decay appear all described by the same Fermi constant. This property is very similar to that of the electromagnetism and the quark-gluon interaction. In either case, once you specify the "charge," the electric charge and the color, respectively, the coupling constants $\alpha$, and $\alpha_s$ describe the strength of the interaction no matter what the particle is. The force does not care the nature of the particle except for its charge. This point strongly suggests that the weak interaction is of the same kind as the other forces, namely the gauge force mediated by the spin one boson similar to the photon and the gluon.

In the modern picture, what causes the nuclear beta decay is the process $d \to u e^- \bar{\nu}_e$. For example, one of the down quarks in the neutron ($udd$) undergoes this decay, converting the neutron to a proton ($udd$). The amplitude for this process is exactly the same as the muon decay with the $V - A$ form.[2]

Let us crudely test the universality based on dimensional analysis. First take the muon decay. All particles in the final state can be approximated massless relative to the muon mass. The decay amplitude is proportional to $G_F$ which has the dimension of energy inverse squared. The decay with is proportional to $G_F^2$. To obtain the decay width that has the dimension of energy, we need to provide a factor with the dimension of energy to the fifth. The only quantity you can use is the muon mass, and hence $\Gamma \sim G_F^2 m_\mu^5$. In fact, much more careful calculation gives

$$\Gamma_\mu = \frac{1}{192\pi^3} G_F^2 m_\mu^5. \tag{4}$$

The factor $1/192\pi^3$ is typical of three-body phase space. Using the measured muon lifetime 2.19703(4) $\mu$s (with higher order corrections), we obtain the Fermi constant $G_F = 1.16639(1) \times 10^{-5} \text{GeV}^2$.

---

[2]Note that the strength is slightly weaker because of the Cabibbo mixing, as we will discuss later.

Table 1: Experimental results ($Q_{EC}$, $t_{1/2}$ and branching ratio, $R$) and calculated correction, $P_{EC}$, for $0^+ \to 0^+$ transitions. Taken from I. S. Towner and J. C. Hardy, `http://www.arxiv.org/abs/nucl-th/9809087`. This table demonstrates that many different nuclei with very different half-lifes $t_{1/2}$ have the same strength of the weak interaction $\mathcal{F}t$ once corrected for the difference in phase space factors.

| | $Q_{EC}$ (keV) | $t_{1/2}$ (ms) | $R$ (%) | $P_{EC}$ (%) | $ft$ (s) | $\mathcal{F}t$ (s) |
|---|---|---|---|---|---|---|
| $^{10}$C | 1907.77(9) | 19290(12) | 1.4645(19) | 0.296 | 3038.7(45) | 3072.9(48) |
| $^{14}$O | 2830.51(22) | 70603(18) | 99.336(10) | 0.087 | 3038.1(18) | 3069.7(26) |
| $^{26m}$Al | 4232.42(35) | 6344.9(19) | $\geq 99.97$ | 0.083 | 3035.8(17) | 3070.0(21) |
| $^{34}$Cl | 5491.71(22) | 1525.76(88) | $\geq 99.988$ | 0.078 | 3048.4(19) | 3070.1(24) |
| $^{38m}$K | 6044.34(12) | 923.95(64) | $\geq 99.998$ | 0.082 | 3049.5(21) | 3071.1(27) |
| $^{42}$Sc | 6425.58(28) | 680.72(26) | 99.9941(14) | 0.095 | 3045.1(14) | 3077.3(23) |
| $^{46}$V | 7050.63(69) | 422.51(11) | 99.9848(13) | 0.096 | 3044.6(18) | 3074.4(27) |
| $^{50}$Mn | 7632.39(28) | 283.25(14) | 99.942(3) | 0.100 | 3043.7(16) | 3073.8(27) |
| $^{54}$Co | 8242.56(28) | 193.270(63) | 99.9955(6) | 0.104 | 3045.8(11) | 3072.2(27) |
| | | | | | Average, $\overline{\mathcal{F}t}$ | 3072.3(9) |
| | | | | | $\chi^2/\nu$ | 1.10 |

The pion lifetime is then estimated as $\Gamma_\pi \sim \frac{1}{8\pi} G_F^2 m_\pi^5$ using the same argument. The factor $1/8\pi$ is typical of two-body phase space. This estimate works quite well to explain the observed lifetime $2.6033 \times 10^{-8}$ s.

What about the neutron lifetime $\tau_n = 885.7(8)$ sec? In this case the factor of energy to the fifth comes from the size of the available phase space. Given $m_n - m_p = 1.29$ MeV, and you have to produce $m_e = 0.511$ MeV, the available energy is less than 1 MeV. Therefore the rough estimate of the decay width is $G_F^2 (1 \text{ MeV})^5/192\pi^3$, and hence the lifetime is roughly $(m_\mu/1 \text{ MeV})^5 \simeq 10^{10}$ longer than the muon, and is of the order of $10^4$ sec. Given how rough this estimate is, this is a reasonable agreement.

Read Cahn–Goldhaber on the two-neutrino experiment which demonstrated that $\nu_e$ and $\nu_\mu$ are indeed distinct particles. The conservation of electron and muon numbers was proposed by Sakata and Inoue in 40's to understand the absence of $\mu \to e\gamma$ decay. It would necessarily require that

two neutrinos have distinct electron and muon numbers and hence must be different particles. The accelerator-based neutrino beam was first produced for this experiment. They brought in a ship wreck as a steel shield to avoid any muons sneaking into the detector from the beam!

# 3   Cabibbo Angle

OK, the Fermi interaction together with the $V - A$ theory seems to describe many weak interactions well. But what about the decay of the strange particles? After all, the strangeness was introduced to explain the pair production of new types of particles via the strong interaction, but their longevity (compared to $10^{-23}$ sec, a characteristic time scale of hadron resonances). The weak interaction is supposed to violate strangeness and let strange particles decay. In modern language, we need the decay of the strange quark $s \rightarrow ue^-\bar{\nu}_e$, very analogous to the nuclear beta decay that is caused by $d \rightarrow ue^-\bar{\nu}_e$.

If you look at the lifetime of $K^+$, $\Lambda$, etc, the universal strength of the weak interaction predicts too short lifetimes. The strange particles are long-lived even in the standard of the weak interaction! However, given the success of the universality in nuclear beta decay and the muon decay, we do not want to give up the idea of universality.

Cabibbo in 1960 proposed a slightly extended notion of universality. He proposed that the weak interaction acts on the linear combination of the down and strange quarks (he didn't use the language of quarks back then; this is the modern translation of what he said),

$$d' = d\cos\theta_C + s\sin\theta_C. \tag{5}$$

The strength of the weak interaction on $d'$ is still the same as in the muon decay. But if you specialize to the $d$, the vertex is smaller by a factor of $\cos\theta_C$, which is still close to one. However, the vertex for the strange quark is much more suppressed, by a factor of $\sin\theta_C \simeq 0.22$. This makes the lifetime of the strange particles longer by a factor of $1/\sin^2\theta_C \simeq 20$.

At this point, the fact that $\cos\theta_C < 1$ could not be established because of the experimental uncertainties. Nowadays, however, we know that indeed the nuclear $\beta$-decay is slightly weaker than the muon decay. In fact, the best test of universality is done using the nuclear beta decay. It gives that the nuclear beta decay is slightly weaker than the muon decay by

a factor of $\cos\theta_C = 0.9740 \pm 0.0005$. On the other hand the decay of strange particles, especially $K^+ \to \pi^0 e^+ \nu_e$ and $K^0 \to \pi^- e^+ \nu_e$, determine $\sin\theta_C = 0.2196 \pm 0.0026$. Putting them together, we find the *total* strength of the quark weak interaction $\sin^2\theta_C + \cos^2\theta_C = 0.9969 \pm 0.0015$. It is consistent with universality (one) at the two sigma level.

# 4   CP Violation

The fact that neutrinos are all left-handed and anti-neutrinos right-handed explicitly break parity, but also charge conjugation. But the product CP still appears to be a symmetry. Under parity, a left-handed neutrino becomes a right-handed neutrino, a state that does not exist. Under charge conjugation, a right-handed neutrino becomes a right-handed anti-neutrino, a state that does exist. Therefore it was hoped that CP is still a symmetry of nature.

It is believed that the combination CPT is an exact symmetry. The CPT theorem states that it is a symmetry to perform charge conjugation C, parity P, and time reversal T at the same time. The assumptions to prove this theorem is very reasonable: (1) Lorentz invariance, (2) hermiticity of the Hamiltonian, and (3) locality. The third one means there is no "action at a distance," but all forces must be mediated by virtual particles which are produced where the sources are. The CPT theorem implies that the particle and anti-particle must have the same mass and the same lifetime (if unstable).

Despite the hope that CP may be preserved by the weak interaction, it fell also shortly after parity violation was discovered. The violation of CP is a fascinating subject, as it is a necessary condition for us to understand why there exists only matter in our Universe but no anti-matter.

## 4.1   Neutral Kaon System

Neutral kaon system is a very strange system, well beyond the fact that they are strange particles anyway. There are two neutral kaons, $K^0(d\bar{s})$ and $\overline{K}^0(s\bar{d})$, anti-particle to each other. Because the strangeness is violated in weak decays, both of them can decay into common final states, $\pi^+\pi^-$ and $\pi^0\pi^0$. Therefore, at the second order in the weak interaction, $K^0$ can go to $\overline{K}^0$. A particle and its anti-particle can mix! Read Cahn–Goldhaber Chapter 7 for how this mixing was proposed, seen and studied.

For quantitative discussion, the Hamitonian for this two-state system at rest can be written as

$$H = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} = \begin{pmatrix} m & \frac{1}{2}\Delta m e^{i\xi} \\ \frac{1}{2}\Delta m e^{-i\xi} & m \end{pmatrix}. \tag{6}$$

(I ignore the factor of $c^2$ here and hereafter.) At the last step, we used the prediction of the CPT theorem that a particle and its anti-particle must have the same mass. The 12 and 21 elements must be complex conjugate to each other because of the hermiticity of the Hamiltonian. The off-diagonal elements represent the mixing between $K^0$ and $\overline{K}^0$.

One important point is that these particles are unstable and decay. To fully take the decay into account, one must include $\pi^+\pi^-$, $\pi^0\pi^0$, $\pi^+\pi^-\pi^0$, $3\pi^0$, $\pi^\pm e^\mp \nu_e$, $\pi^\pm \mu^\mp \nu_\mu$ etc states into the Hilbert space and the Hamiltonian becomes an infinite-dimensional matrix. That would make the discussion intractable. Fortunately, there is a simple way to account for the decay into multi-pion states, just by allowing the Hamiltonian to be non-hermitian. We include the exponential decay as a part of the Hamiltonian. For example, an unstable particle of mass $m$ and width $\Gamma$ would have a time-evolution $e^{-i(m-i\Gamma/2)t}$ such that the probability goes down exponentially $|e^{-i(m-i\Gamma/2)t}|^2 = e^{-\Gamma t}$. (I ignored the factor of $\hbar$.) We can regard $m - i\frac{1}{2}\Gamma$ as the eigenvalue of the Hamiltonian. The imaginary part (anti-hermitian one-by-one matrix) represents the decay. Similarly in our case of two-state sytem, we include an anti-hermitian piece

$$
\begin{aligned}
H &= \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} - i\frac{1}{2}\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \\
&= \begin{pmatrix} m & \frac{1}{2}\Delta m e^{i\xi} \\ \frac{1}{2}\Delta m e^{-i\xi} & m \end{pmatrix} - i\frac{1}{2}\begin{pmatrix} \Gamma & \frac{1}{2}\Delta\Gamma e^{i\gamma} \\ \frac{1}{2}\Delta\Gamma e^{-i\gamma} & \Gamma \end{pmatrix} \tag{7}
\end{aligned}
$$

Again we used the CPT relation $\Gamma_{11} = \Gamma_{22}$ and anti-hermiticity of the second term $\Gamma_{12} = \Gamma_{21}^*$.

Even though the matrix is no longer hermitian, it is nonetheless a simple two-by-two matrix and can be diagonalized. The eigenstates are given by

$$|K_S\rangle = p|K^0\rangle + q|\overline{K}^0\rangle, \qquad |K_L\rangle = p|K^0\rangle - q|\overline{K}^0\rangle, \tag{8}$$

where the complex coefficients $p$, $q$ are normalized $|p|^2 + |q|^2 = 1$ and

$$\frac{q}{p} = +\sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{9}$$

The notation refers to "K-short" and "K-long" because their lifetimes are very different as we will see below.

Note that CP interchanges $K^0$ and $\overline{K}^0$, and hence the 12 and 21 elements. Therefore, the complex $M_{12}$ and $\Gamma_{12}$ break CP invariance. We will see below the difference between the real and complex cases.

### 4.1.1 CP Conserving Case

Let us suppose for the moment that $M_{12}$, $\Gamma_{12}$ are real. This corresponds to the case where CP is conserved. In this case, $p = q$ and the two eigenstates are

$$|K_1\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\overline{K}^0\rangle), \qquad |K_2\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\overline{K}^0\rangle). \qquad (10)$$

The eigenvalues are

$$H|K_1\rangle = (m_1 - i\frac{1}{2}\Gamma_1)|K_1\rangle = (m + \frac{\Delta m}{2} - \frac{i}{2}\Gamma - \frac{i}{4}\Delta\Gamma)|K_1\rangle, \quad (11)$$

$$H|K_2\rangle = (m_2 - i\frac{1}{2}\Gamma_2)|K_2\rangle = (m - \frac{\Delta m}{2} - \frac{i}{2}\Gamma + \frac{i}{4}\Delta\Gamma)|K_2\rangle. \quad (12)$$

The notation $\Delta m$ and $\Delta\Gamma$ therefore refers to the difference in the mass and the width between two states.

The convention commonly used in the literature is that the charge conjugation interchanges $K^0$ and $\overline{K}^0$ with a minus sign (I don't know why),

$$C|K^0\rangle = -|\overline{K}^0\rangle, \qquad C|\overline{K}^0\rangle = -|K^0\rangle. \qquad (13)$$

Because they are pseudoscalar $0^-$ mesons, the parity changes their signs, too,

$$P|K^0\rangle = -|K^0\rangle, \qquad P|\overline{K}^0\rangle = -|\overline{K}^0\rangle. \qquad (14)$$

Putting them together, CP interchanges them without an additional sign,

$$CP|K^0\rangle = +|\overline{K}^0\rangle, \qquad CP|\overline{K}^0\rangle = +|K^0\rangle. \qquad (15)$$

The states $|K_{1,2}\rangle$ therefore are CP eigenstates,

$$CP|K_1\rangle = +|K_1\rangle, \qquad CP|K_2\rangle = -|K_2\rangle. \qquad (16)$$

When they decay, for example into $\pi^0\pi^0$, the two pions must be in $L = 0$ state because neither the initial state particle (kaon) nor the final state

particles (pions) have spins and the angular momentum must be conserved. Therefore the parity of the final state is $(-1)^L(-1)^2 = +1$, where two factors of $(-1)$ come from the intrinsic parity of the pion. Under charge conjugation, $\pi^0$ is an eigenstate of $C = +1$ (remember it decays into two photons and we could also derive this eigenvalue from the quark model) and hence $\pi^0\pi^0$ state is also a state with $C = +1$. Therefore, the CP eigenvalue of the final state is $CP = +1$. If CP is conserved, only $K_1$ can decay into this final state.

On the other hand, $K_2$ has odd CP eigenvalue, and (if CP is conserved) must decay into a state with odd CP eigenvalue. $3\pi^0$ and $\pi^+\pi^-\pi^0$ state have odd CP and the decay of $K_2$ into these final states are allowed. However such a decay is suppressed because of much smaller phase space. The kaon mass is $497.672\pm0.031$ MeV, the charged pion mass $139.57018\pm0.00035$ MeV and the neutral pion mass $134.9766\pm0.0006$ MeV. The sum of masses for $\pi^+\pi^-\pi^0$ state is $414.117$ MeV, quite close to the mass of the kaon. Therefore this decay is barely allowed, and it makes the lifetime of $K_2$ much longer than that of $K_1$. In fact, the lifetimes are $0.8935(8) \times 10^{-10}$ sec and $5.17(4) \times 10^{-8}$ sec, respectively, with almost three orders of magnitude difference.

Suppose $K^0$ is produced in a collision of a proton on a nucleus target, with the final state $pp \to K^0\Sigma^+p$. Because the strangeness is conserved in the strong interaction, the existence of $\Sigma^+$ implies that the kaon must be $K^0$, not $\overline{K}^0$. Once it is produced, however, $K^0$ starts to mix with $\overline{K}^0$. The time evolution is given simply by the Schrödinger equation, with the phase factor $e^{-iEt}$ on Hamiltonian eigenstates. Therefore,

$$|K^0(t)\rangle = e^{-iHt}\frac{1}{\sqrt{2}}(|K_1\rangle+|K_2\rangle) = \frac{1}{\sqrt{2}}(|K_1\rangle e^{-i(m_1-i\Gamma_1/2)t}+|K_2\rangle e^{-i(m_2-i\Gamma_2/2)t}).$$
(17)

Given this we can calculate the probability of finding $K^0$ or $\overline{K}^0$ at a given moment,

$$
\begin{aligned}
P(K^0 \to K^0, t) &= |\langle K^0|K^0(t)\rangle|^2 = \left|\frac{1}{2}\left(e^{-i(m_1-i\Gamma_1/2)t} + e^{-i(m_2-i\Gamma_2/2)t}\right)\right|^2 \\
&= \frac{1}{4}\left(e^{-\Gamma_1 t} + e^{-\Gamma_2 t} + 2(\cos\Delta mt)e^{-(\Gamma_1+\Gamma_2)t/2}\right), \quad (18)
\end{aligned}
$$

$$
\begin{aligned}
P(K^0 \to \overline{K}^0, t) &= |\langle \overline{K}^0|K^0(t)\rangle|^2 = \left|\frac{1}{2}\left(e^{-i(m_1-i\Gamma_1/2)t} - e^{-i(m_2-i\Gamma_2/2)t}\right)\right|^2 \\
&= \frac{1}{4}\left(e^{-\Gamma_1 t} + e^{-\Gamma_2 t} - 2(\cos\Delta mt)e^{-(\Gamma_1+\Gamma_2)t/2}\right), \quad (19)
\end{aligned}
$$

The last term comes from the interference of two Hamiltonian eigenstates

with $\Delta m = m_1 - m_2$. When the kaon is moving relativistically (as is often the case in kaon experiments), of course the time-dilation effect must be considered as well, so that $t$ is replaced by $\gamma t$.

What you see here is that neutral kaons *oscillate*. If you ignore the decay, the initial state $K^0$ can evolve to its anti-particle, a pure $\overline{K}^0$ state, when $t = \pi/\Delta m$, and comes back to a pure $K^0$. In real life, the oscillation frequency is close to the lifetime of the $K_S$, $\Delta m = 0.5303(0) \times 10^{10}$ sec$^{-1}$. Therefore over a few oscillations practically all $K_S$ will be gone.

CPLEAR experiment has the best measurement of $\Delta m$ so far. They collided beams of protons and anti-protons at very low energies. They of course annihilate when they meet, and can produce $K^0 K^- \pi^+$ or $\overline{K}^0 K^+ \pi^-$ final states. Because of the conserved strangeness in the strong interaction, tagging $K^-$ or $K^+$ in the final state determines if the neutral kaon is $K^0$ or $\overline{K}^0$ at the time of production. After they are produced, however, they mix in the course of time evolution, and decay. If you detect the decay into $\pi^\pm e^\mp \nu_e$, it can tell you if the particle that decayed was $K^0$ or $\overline{K}^0$. This way, you can measure the probabilities of the produced $K^0$ to decay as $\overline{K}^0$ or vice versa. They plotted the asymmetry

$$
\begin{aligned}
A_{\Delta m} &= \frac{R(K^0 \to e^+ \pi^- \nu_e) - R(K^0 \to e^- \pi^+ \bar\nu_e)}{R(K^0 \to e^+ \pi^- \nu_e) + R(K^0 \to e^- \pi^+ \bar\nu_e)} \\
&= \frac{2(\cos \Delta m t)e^{-(\Gamma_S + \Gamma_L)t/2}}{e^{-\Gamma_S t} + e^{-\Gamma_L t}}.
\end{aligned}
\tag{20}
$$

The data Fig. 2 beautifully agree with the expression.

### 4.1.2   Indirect CP Violation

The real surpise came in 1964 when Fitch and Cronin followed the decay of neutral kaons after waiting a long time (actually over a long distance, rather, as kaons were moving at relativistic speeds). Because the lifetimes of two eigenstates are so different, you can be quite sure that you don't have any shorter-lived state anymore. Yet, they had observed decay of the neutral kaon into $\pi^+ \pi^-$. This means that the state with longer lifetime has a small mixture of CP even component. In other words,

$$
|K_L\rangle = |K_2\rangle + \epsilon |K_1\rangle.
\tag{21}
$$

It is not an eigenstate of CP anymore. This is the discovery of CP violation. The weak interaction violates not only C and P but also CP!
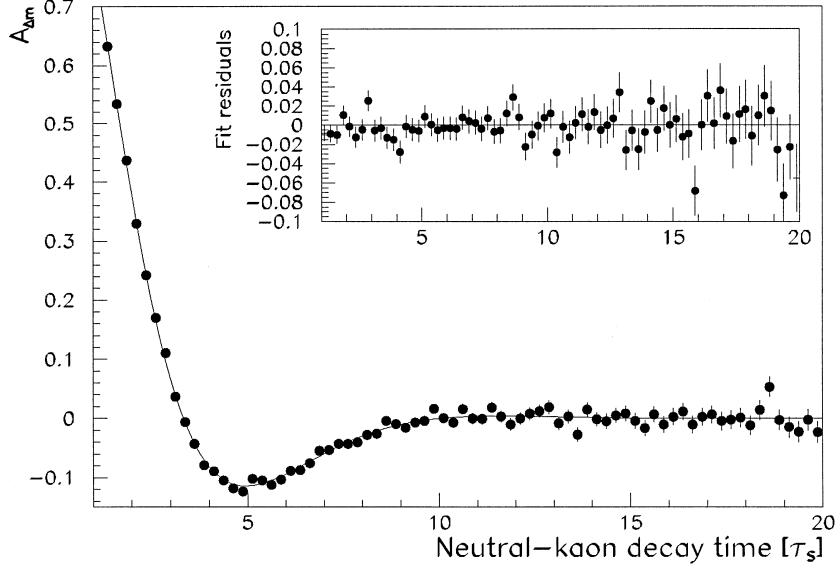
Figure 2: CPLEAR measurement of the asymmetry in $K^0$-$\overline{K}^0$ oscillation. Taken from *Phys. Lett.* **B 444**, 38 (1998).

In order to explain this small admixture of the "wrong" CP state in the Hamiltonian eigenstate, we go back to the general case of complex $M_{12}$ and $\Gamma_{12}$. The eigenstates are given by

$$|K_S\rangle = p|K^0\rangle + q|\overline{K}^0\rangle, \qquad |K_L\rangle = p|K^0\rangle - q|\overline{K}^0\rangle, \qquad (22)$$

with the complex coefficients $p$, $q$

$$\frac{q}{p} = +\sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \qquad (23)$$

Note that the phase of $M_{12}$ and $\Gamma_{12}$ can be changed simultaneously by re-defining the phase of states $K^0$ and $\overline{K}^0$, but the relative phase between $M_{12}$ and $\Gamma_{12}$ is convention independent. For instance, we can always choose our convention that $\Gamma_{12}$ is real. Then $M_{12}$ has an imaginary part $M_{12} = \Re e(M_{12}) + i\Im m(M_{12})$. Assume $\Im m(M_{12}) \ll \Re e(M_{12})$ because it turns out to be true. Then,

$$\frac{q}{p} = +\sqrt{\frac{\Re e(M_{12}) - i\Im m(M_{12}) - \frac{i}{2}\Gamma_{12}}{\Re e(M_{12}) + i\Im m(M_{12}) - \frac{i}{2}\Gamma_{12}}} \approx 1 + \frac{-i\Im m(M_{12})}{\Re e(M_{12}) - \frac{i}{2}\Gamma_{12}} \equiv 1 - 2\epsilon. \quad (24)$$

16

With this definition of $\epsilon$, we find $q = \frac{1}{\sqrt{2}}(1 - \epsilon)$, $p = \frac{1}{\sqrt{2}}(1 + \epsilon)$ (all quantities valid only up to the first order in $\epsilon$), and

$$|K_L\rangle = p|K^0\rangle - q|\overline{K}^0\rangle = \frac{1}{\sqrt{2}}(1+\epsilon)|K^0\rangle - \frac{1}{\sqrt{2}}(1-\epsilon)|\overline{K}^0\rangle = |K_2\rangle + \epsilon|K_1\rangle, \quad (25)$$

as desired. The data on $K_L \to \pi\pi$ determines the magnitude of $\epsilon$ as $|\epsilon| = 2.287(17) \times 10^{-3}$ a very small number but non-vanishing.

Another consequence of $\epsilon \neq 0$ is that $K_L \to \pi^- e^+ \nu_e$ and $K_L \to \pi^+ e^- \bar{\nu}_e$ have different rates. For these decays, $K^0$ or $\overline{K}^0$ are singled out. For instance, $K^0(d\bar{s})$ can decay into $\pi^- e^+ \nu_e$ because $\bar{s} \to \bar{u} e^- \bar{\nu}_e$, but cannot decay into $\pi^+ e^- \bar{\nu}_e$. The same is true for muon final states. Therefore, the decays into these states "measure" $|q|/|p|$ directly. People define the observable

$$\delta_L = \frac{\Gamma(K_L \to \pi^- \ell^+ \nu_\ell) - \Gamma(K_L \to \pi^+ \ell^- \bar{\nu}_\ell))}{\Gamma(K_L \to \pi^- \ell^+ \nu_\ell) + \Gamma(K_L \to \pi^+ \ell^- \bar{\nu}_\ell))}, \quad (26)$$

where $\ell = e, \mu$. Experimentally, $\delta_L = 3.27(12) \times 10^{-3}$. It follows from the above discussions that $\delta_L = 2\Re e(\epsilon)$. Using the observed values of $\Delta m$ and $|\Gamma_{12}| = \Gamma_S - \Gamma_L$, you can verify that this observed value of $\delta_L$ is consistent with $|\epsilon|$ mentioned above.

Therefore, matter and anti-matter mix, but not quite 50:50. The world of anti-matter is not an exact mirror of the world of matter. This phenomenon is called "indirect CP Violation" because the violation of CP symmetry is in the mixing $\Im m(M_{12}) \neq 0$ but not in the decay of the particles.

### 4.1.3  Direct CP Violation

Much more recently, even more direct but subtle difference between $K^0$ and $\overline{K}^0$ became known. This phenomenon is called "direct CP violation," because it is the difference in the decays or particle and its anti-particle counterpart. This phenomenon was established by 2001.

The direction CP violation is the difference

$$\Gamma(K^0 \to \pi^+ \pi^-) - \Gamma(\overline{K}^0 \to \pi^+ \pi^-), \quad (27)$$

and similarly for $\pi^0 \pi^0$. Because of the CPT theorem, the total width $\Gamma(K^0 \to \text{everything})$ and $\Gamma(\overline{K}^0 \to \text{everything})$ are equal. However the individual contributions can be different if CP is violated. This difference is extremely

small, at the level of one part in million, but had been observed by KTeV and NA48 collaborations after heroic efforts.

The measured quantity is what is called $\epsilon'$, defined by

$$\eta_{+-} = \frac{A(K_L \to \pi^+\pi^-)}{A(K_S \to \pi^+\pi^-)} \approx \epsilon + \epsilon', \tag{28}$$

$$\eta_0 = \frac{A(K_L \to \pi^0\pi^0)}{A(K_S \to \pi^0\pi^0)} \approx \epsilon - 2\epsilon'. \tag{29}$$

$$\tag{30}$$

$\epsilon'$ can be measured in the double ratio,

$$\frac{|\eta_{00}|^2}{|\eta_{+-}|^2} \approx 1 - 6\Re e\frac{\epsilon'}{\epsilon}. \tag{31}$$

If you go into the details of the mixing in the $K^0$ and $\overline{K}^0$ system, this quantity can be seen as the direct CP violation.

### 4.1.4   T Violation

Because CPT must be a true symmetry of nature, according to the Quantum Field Theory, CP violation implies T violation. The evidence for T violation was found by CPLEAR experiment. They collided beams of protons and anti-protons at very low energies. They of course annihilate when they meet, and can produce $K^0 K^- \pi^+$ or $\overline{K}^0 K^+ \pi^-$ final states. Because of the conserved strangeness in the strong interaction, tagging $K^-$ or $K^+$ in the final state determines if the neutral kaon is $K^0$ or $\overline{K}^0$ at the time of production. After they are produced, however, they mix in the course of time evolution, and decay. If you detect the decay into $\pi^\pm e^\mp \nu_e$, it can tell you if the particle that decayed was $K^0$ or $\overline{K}^0$. This way, you can measure the probabilities of the produced $K^0$ to decay as $\overline{K}^0$ or vice versa. They have found a difference in the probability of $K^0 \to \overline{K}^0$ and $\overline{K}^0 \to K^0$ consistent with the level expected. They reported

$$\left\langle \frac{R(\overline{K}^0(0) \to e^+\pi^-\nu_e(\tau)) - R(K^0(0) \to e^-\pi^+\bar\nu_e(\tau))}{R(\overline{K}^0(0) \to e^+\pi^-\nu_e(\tau)) + R(K^0(0) \to e^-\pi^+\bar\nu_e(\tau))} \right\rangle = (6.6\pm1.3\pm1.0)\times10^{-3},$$

$$\tag{32}$$

where $\langle...\rangle$ is an average over the time window $1\tau_S < \tau < 20\tau_S$ ($\tau_S$ is the lifetime of $K_S$).

Theoretically, we start with

$$|K^0\rangle = \frac{1}{2p}(|K_S\rangle + |K_L\rangle). \tag{33}$$

The time evolution gives

$$
\begin{aligned}
|K^0(t)\rangle &= \frac{1}{2p}(|K_S\rangle e^{-im_S t - \Gamma_S t/s} + |K_L\rangle e^{-im_L t - \Gamma_L t/s}) \\
&= \frac{1}{2p}((p|K^0\rangle + q|\overline{K}^0\rangle)e^{-im_S t - \Gamma_S t/s} + (p|K^0\rangle - q|\overline{K}^0\rangle)e^{-im_L t - \Gamma_L t/s}) \\
&= |K^0\rangle\frac{1}{2}(e^{-im_S t - \Gamma_S t/s} + e^{-im_L t - \Gamma_L t/s}) \\
&\quad + |\overline{K}^0\rangle\frac{q}{2p}(e^{-im_S t - \Gamma_S t/s} - e^{-im_L t - \Gamma_L t/s}). \tag{34}
\end{aligned}
$$

Therefore, for $K^0$ produced at $t = 0$ to decay as $\overline{K}^0$ at $\tau$ is

$$R(K^0(0) \to e^-\pi^+\bar{\nu}_e(\tau))\frac{|q|^2}{4|p|^2}(e^{-\Gamma_S \tau} + e^{-\Gamma_L \tau} - (\cos\Delta m\tau)e^{-(\Gamma_S + \Gamma_L)\tau/2}). \tag{35}$$

Similarly,

$$R(\overline{K}^0(0) \to e^+\pi^-\nu_e(\tau))\frac{|p|^2}{4|q|^2}(e^{-\Gamma_S \tau} + e^{-\Gamma_L \tau} - (\cos\Delta m\tau)e^{-(\Gamma_S + \Gamma_L)\tau/2}). \tag{36}$$

Therefore,

$$\frac{R(\overline{K}^0(0) \to e^+\pi^-\nu_e(\tau)) - R(K^0(0) \to e^-\pi^+\bar{\nu}_e(\tau))}{R(\overline{K}^0(0) \to e^+\pi^-\nu_e(\tau)) + R(K^0(0) \to e^-\pi^+\bar{\nu}_e(\tau))} = \frac{|p|^4 - |q|^4}{|p|^4 + |q|^4} = 4\Re e(\epsilon). \tag{37}$$

Indeed, the observed T-violation is consistent with this expectation.

## 4.2   Neutral $B$-Meson System

The formalism used in the neutral kaon system can be used also in the neutral $B$-meson system, $B^0(d\bar{b})$ and $\overline{B}^0(b\bar{d})$. The main difference in this case is that the lifetime difference is very small. Therefore, the observation of CP violation would use a different technique.

The idea is very simple. You first produce $\Upsilon(4S)$ resonance in an $e^+e^-$ annihilation, which decays into $B^0$-$\overline{B}^0$ pair. The point is that $\Upsilon$ is a spin one

19

resonance, and hence there must be $L = 1$ angular momentum in the final state ($B^0$ is a pseudo-scalar mesons just like the pion). But $L = 1$ means that the interchange of two particles gives $-1$, which is not possible for identical bosons. Therefore, even though $B^0$ and $\overline{B}^0$ states evolve in time, you can never have a moment when both of them are simultaneously $B^0$ (or $\overline{B}^0$). In other words, if you are sure that one of them is $B^0$ at a given moment $t$, the other particle must be $\overline{B}^0$ at the same moment $t$. This way, you can identify a particle to be 100% $B^0$ at the moment $t$, and see how it evolves in time.

If you look for a specific final state such as $J/\psi K_S$, it is a CP eigenstate ($CP = -1$ to the extent we ignore $\epsilon$ mixture of CP-odd state in $K_S$) and both $B^0$ and $\overline{B}^0$ can decay into this state. This final state is particularly easy to identify when $J/\psi \rightarrow \mu^+\mu^-$ or $e^+e^-$ and $K_S \rightarrow \pi^+\pi^-$ with the well-known masses. This way, you can define a time-dependent asymmetry,

$$a_f(t) = \frac{\Gamma(\overline{B}^0(t) \rightarrow f) - \Gamma(B^0(t) \rightarrow f)}{\Gamma(\overline{B}^0(t) \rightarrow f) + \Gamma(B^0(t) \rightarrow f)} \tag{38}$$

A problem to measure this quantity is that you need to resolve "time" when $B$-meson decays. Pier Oddone at Lawrence Berkeley Laboratory came up with the idea that you can collide two beams asymmetrically, so that the produce $B$ and $\overline{B}$ mesons move in the laboratory frame. Then you can determine the vertex where they decay and translate it to the timing. This way, you can see if there is any time-dependent difference between the $B^0$ and $\overline{B}^0$ decay.

The expectation is easily derived by using the same type of equation as in the kaon case,

$$|B_1\rangle = p|B^0\rangle + q|\overline{B}^0\rangle, \qquad |B_2\rangle = p|B^0\rangle - q|\overline{B}^0\rangle, \tag{39}$$

where the complex coefficients $p$, $q$ are normalized $|p|^2 + |q|^2 = 1$ and

$$\frac{q}{p} = +\sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{40}$$

We don't use "short" or "long" because the lifetime difference is small. (Unlike in the case of kaons, where $K_{1,2}$ were reserved for CP eigenstate, $B_{1,2}$ are not CP eigenstates but rather Hamiltonian eigenstates.) When you start with $B^0$ at $t = 0$,

$$|B^0(0)\rangle = \frac{1}{2p}(|B_1\rangle + |B_2\rangle), \tag{41}$$

20

its time evolution is given by

$$|B^0(t)\rangle = \frac{1}{2p}(|B_1\rangle e^{-im_1t-\Gamma_1t/2} + |B_2\rangle e^{-im_2t-\Gamma_2t/2})$$
$$= |B^0\rangle(e^{-im_1t-\Gamma_1t/2} + e^{-im_2t-\Gamma_2t/2}) + \frac{q}{p}|\overline{B}^0\rangle(e^{-im_1t-\Gamma_1t/2} - e^{-im_2t-\Gamma_2t/2}).$$
$$(42)$$

Now we let this particle decay into a CP eigenstate $CP|f\rangle = \eta_f|f\rangle$ with $\eta_f = \pm 1$. We assume that there is no direct CP violation, namely that decay amplitudes are CP invariant, and theonly CP violation is in the mixing $q/p \neq 1$. Then, the decay amplitudes

$$A = \langle f|B^0\rangle, \qquad \bar{A} = \langle f|\overline{B}^0\rangle \tag{43}$$

are related by

$$\bar{A} = \langle f|\overline{B}^0\rangle = \langle f|CP|B^0\rangle = \eta_f\langle f|B^0\rangle = \eta_f A. \tag{44}$$

In the case of $B^0, \overline{B}^0 \rightarrow J/\psi K_S$, the CP symmetry is good.[3] In general, we define the ratio

$$\bar{\rho}(f) = \frac{\bar{A}}{A}. \tag{45}$$

For the decay process of our interest, $\bar{\rho}(J/\psi K_S) = 1$.

Then the decay amplitude for $B^0$ to decay into the state $f$ after time interval $t$ is

$$\langle f|B^0(t)\rangle = A(e^{-im_1t-\Gamma_1t/2} + e^{-im_2t-\Gamma_2t/2}) + \frac{q}{p}\bar{A}(e^{-im_1t-\Gamma_1t/2} - e^{-im_2t-\Gamma_2t/2})$$
$$= A\left((e^{-im_1t} + e^{-im_2t}) + \eta_f\frac{q}{p}\bar{\rho}(f)(e^{-im_1t} - e^{-im_2t})\right)e^{-\Gamma t/2}. \quad (46)$$

Here, I used the approximation that the lifetime difference is small $\Gamma_1 \approx \Gamma_2$. The same quantity for $\overline{B}^0$ can be worked analogously and we find

$$\langle f|\overline{B}^0(t)\rangle = \bar{A}\left(\eta_f(e^{-im_1t} + e^{-im_2t}) + \frac{p}{q}\frac{1}{\bar{\rho}(f)}(e^{-im_1t} - e^{-im_2t})\right)e^{-\Gamma t/2}. \quad (47)$$

---

[3]I've ignored the so-called "strong phase" in this discussion, which allows different phases for $A$ and $\bar{A}$ but $|A| = |\bar{A}|$. The strong phase is also small in this final state.

Using the above results, we can easily work out the time-dependent asymmetry,

$$a_f(t) = \eta_f \Im m \left( \frac{q}{p} \frac{\bar{A}}{A} \right) \sin \Delta m t. \tag{48}$$

Here is it assumed that there is no direct CP violation, namely $|\frac{q}{p} \frac{\bar{A}}{A}| = 1$. The Belle and BABAR collaborations have observed this asymmetry beautifully in 2002, and agree with each other.
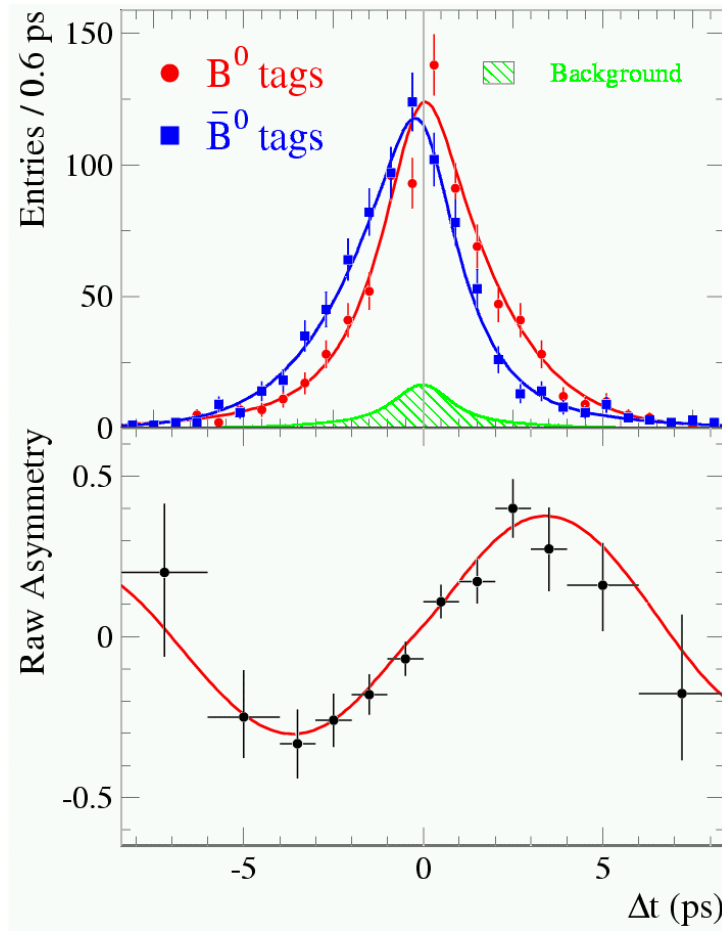


Figure 3: Time-dependent asymmetry in $B^0 \to J/\psi K_S$ decay from BABAR experiment at SLAC.

# 5  Glashow–Weinberg–Salam Theory

We have seen that the weak interaction happens "pretty much" on the isospin lowering or raising operator, except for small admixture of the strange quark. On the side of leptons, they are strictly between $e$ and $\nu_e$, or $\mu$ and $\nu_\mu$. Moverover, the strength of the weak interaction is universal, *i.e.*, it is described by the same Fermi constant for any processes. This point strongly suggests that the weak interaction is caused by the similar mechanism as the electromagnetism or quark-gluon theory of the strong interaction, where the strength of the force is determined by the overall coupling constant and the charges (or matrices) of given particle type. Based on this motivation, Sheldon Glashow in 1962, and later Steven Weinberg and Abdus Salam, tried to formulate the theory of the weak interaction in terms of the same type of theory, namely gauge theory.

Recall the case of the gluon. It is based on three colors, and you demand that you can freely rotate among three colors. This defines three-by-three unitarity rotations, given by the group $SU(3)$ (after removing the overall phase part). There are eight generators for this group, given by Gell-Mann's lambda matrices (with a factor of 1/2). There are eight gluons, each of which couples to each generator. When the generator acts on a quark, it may change its color. In the same way, we look at the "weak isospin" $SU(2)$. We distinguish it from the "strong isospin" which is between up and down with no admixture of strange, and has nothing to do with leptons. We put together doublets

$$
\begin{pmatrix} u \\ d' \end{pmatrix}, \qquad \begin{pmatrix} \nu_e \\ e \end{pmatrix}, \qquad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}. \tag{49}
$$

Under the two-by-two unitarity rotation without the overall phase, the $SU(2)$ group, there are three generators given by the three Pauli matrices $\tau_1$, $\tau_2$, $\tau_3$ (with a factor of 1/2). Accordingly, we postulate three $W$-bosons, $W_1$, $W_2$, $W_3$. Obviously $W$ stands for "weak." When they interact with doublets, they produce a set of amplitudes given by

$$
(\bar{\nu}_\mu, \bar{\mu})(g\vec{W} \cdot \frac{\vec{\tau}}{2}) \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} = (\bar{\nu}_\mu, \bar{\mu})\frac{g}{2} \begin{pmatrix} W_3 & W_1 - iW_2 \\ W_1 + iW_2 & -W_3 \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}. \tag{50}
$$

We introduce the notation $W^+ = (W_1 - iW_2)/\sqrt{2}$, $W^- = (W_1 + iW_2)/\sqrt{2}$,

so that it becomes

$$(\bar{\nu}_\mu, \bar{\mu})\frac{g}{2} \begin{pmatrix} W_3 & \sqrt{2}W^+ \\ \sqrt{2}W^- & -W_3 \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}. \tag{51}$$

The same set is there also for other doublets. Therefore, there are vertices such as $\mu^- \to \nu_\mu W^-$, $\nu_e \to e^- W^+$, etc. All these vertices come with the universal strength $g$, analogous to $e$ in the case of electromagnetism.

This is very nice. The muon decay, for example, happens because we use the vertex $\mu^- \to \nu_\mu W^-$, where $W^-$ is virtual. This virtual $W^-$ has to materialize quickly, such as $W^- \to e^- \bar{\nu}_e$. If you work out numbers, the Fermi constant is then given by

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2}. \tag{52}$$

You get two factors of $g$ because you use two vertices, and the amplitude is suppressed by the heaviness of the $W$-boson. The same can be said about the neutron beta decay, where one of the down-quarks in the neutron emits $d \to uW^-$, followed quickly by $W^- \to e^- \bar{\nu}_e$. Because all these are based on the same interaction, no wonder why they come out universal (modulo the mixing of the strange quark). Hurray! We got the theory of the weak force!

But it is a little bit too early to celebrate a victory. There are several tricky points which we haven't seen in other forces we have to deal with.

One is that the weak force is of $V - A$, and hence only the chirality left $\gamma_5 = -1$ component couples to it. The way to achieve it is to assume that the right-handed chirality particles are singlets under the weak isospin, so that $W$-bosons don't "see" them, in the same way that the photons don't see neutrinos because of the lack of electric charge. $W$-bosons couple to "doubletness," and they do not interact with singlets. Therefore, we have to accept that the particles are organized as

$$\begin{pmatrix} u_L \\ d'_L \end{pmatrix}, u_R, d_R, \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}, e_R, \begin{pmatrix} \nu_\mu \\ \mu_L \end{pmatrix}, \mu_R. \tag{53}$$

I didn't put the subscript $L$ on neutrinos because they are all left-handed. It looks very strange, but this is the only way we can explain the $V - A$ nature of the weak interaction. In other words, we treat left-handed and right-handed particles as *different particles*. Of course this raises the concern how come they can be massive, because the chirality is not a good quantum number

24

for massive fermions as we discussed before. We will come back to the issue of the mass later, and let us forge ahead for the moment.

The next tricky point is that we have to somehow get photon out of this. The electric charge does not commute with the isospin raising or lowering operators, so that it must be a part of the group we introduce to explain the forces. The natural candidate here is of course $W_3$. However, we can see immediately that this won't work. $W_3$ couples to neutrinos! We do not want photons to interact with neutrinos because they are electrically neutral. We are still missing something.

What it means is that the $SU(2)$ is not enough to explain the weak interaction and the electromagnetism at the same time. We need something more. And we would like to do so without any more proliferation of particles.

The only way to get out of this dilemma is to introduce another charge, called "weak hypercharge" (this is another bad example of reusing the same name to mean something completely different). We also introduce a new spin one hypercharge boson $B$, similarly as the photon (that is why it is $B$, next to $A$ for the vector potential). To get started we assign the charge $-1/2$ to the lepton doublets. We could have chosen a different charge, but this turns out to be convenient. Then the set of amplitudes we deal with includes the $B$-boson,

$$(\bar{\nu}_\mu, \bar{\mu})(g\vec{W}\cdot\frac{\vec{\tau}}{2}+g'B\frac{-1}{2})\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} = (\bar{\nu}_\mu, \bar{\mu})\frac{1}{2}\begin{pmatrix} gW_3 - g'B & g\sqrt{2}W^- \\ g\sqrt{2}W^+ & -gW_3 - g'B \end{pmatrix}\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}.$$
(54)

We have two neutral spin one bosons, $B$ and $W_3$. No matter what we do, we end up with two neutral spin one bosons, and one of them must be our good-old photon. Then the combination that couples to the neutrino had better be the other one. In other words, there must be a new spin one boson, called $Z$,[4] given by

$$\sqrt{g^2 + g'^2}Z = gW_3 - g'B.$$
(55)

The prefactor is the normalization factor. The photon must be the orthogonal combination,

$$\sqrt{g^2 + g'^2}A = g'W_3 + gB.$$
(56)

Instead of always referring to the coupling constants $g$ and $g'$, it is convenient

---

[4]This is another arrogant name next to $\Omega^-$. This is the "last" boson. Later on, when people started to discuss grand unified theories, we had to call new bosons $X$ and $Y$.

to introduce the "weak mixing angle,"[5]

$$\theta_W = \sin^{-1} \frac{g'}{\sqrt{g^2 + g'^2}}. \tag{57}$$

This allows us to rewrite the $Z$ and photon as simple as

$$Z = W_3 \cos\theta_W - B \sin\theta_W, \tag{58}$$
$$A = W_3 \sin\theta_W + B \cos\theta_W. \tag{59}$$

Not only that we mix quarks, we also have to mix gauge bosons.

Once we have done this, the interaction of the photon and the $Z$-boson to the muon (or electron) is fixed uniquely,

$$\begin{aligned}
\frac{1}{2}(-gW_3 - g'B) &= \frac{1}{2}(-g(Z \cos\theta_W + A \sin\theta_W) - g'(-Z \sin\theta_W + A \cos\theta_W)) \\
&= \frac{1}{2}(-g \cos\theta_W + g' \sin\theta_W)Z - \frac{1}{2}(g \sin\theta_W + g' \cos\theta_W)A.
\end{aligned} \tag{60}$$

The coupling to the photon must be $e = -|e|$,

$$|e| = \frac{1}{2}(g \sin\theta_W + g' \cos\theta_W) = \frac{gg'}{\sqrt{g^2 + g'^2}}. \tag{61}$$

In other words, one combination of two new coupling constants we introduced is already measured, and the unknown parameter is just $\theta_W$,

$$g = \frac{|e|}{\sin\theta_W}, \qquad g' = \frac{|e|}{\cos\theta_W}. \tag{62}$$

Then the coupling of the left-handed muon (or electron) to the $Z$-boson is also determined,

$$\frac{1}{2}(-g \cos\theta_W + g' \sin\theta_W) = \frac{|e|}{\sin\theta_W \cos\theta_W}\left(-\frac{1}{2} + \sin^2\theta_W\right). \tag{63}$$

---

[5]I grew up with people who called it "Weinberg angle." But Weinberg's paper actually does not introduce this angle, but rather uses $g$ and $g'$ throughout. It was Glashow who introduced this angle earlier. But $\theta_W$ is a wide-spread notation and we can't call it Glashow angle anymore. Recently the name "weak mixing angle" is used more widely that solves this problem.

In general, any particle would couple to the combination $gI_3W_3 + g'YB$. Here, $I_3$ is the third component of the isospin $\pm\frac{1}{2}$ for doublets and 0 for singlets. $Y$ is the "weak hypercharge," which we would like to determine now. By rewriting this combination using the $Z$-boson and the photon, we find

$$gI_3(Z\cos\theta_W + A\sin\theta_W) + g'Y(-Z\sin\theta_W + A\cos\theta_W)$$
$$= \frac{|e|}{\sin\theta_W\cos\theta_W}(I_3\cos^2\theta_W - g'Y\sin^2\theta_W)Z + |e|(I_3 + Y)A. \quad (64)$$

The electric charge of the particle is then given by $Q = I_3 + Y$. This helps us to determine what we have to take for hypercharges for individual particles. This is the "weak" analogue of Gell-Mann–Nishijima relation. The only possible hypercharge assignement is

$$\begin{pmatrix} u_L \\ d'_L \end{pmatrix}^{+1/6}, u_R^{+2/3}, d_R^{-1/3}, \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}^{-1/2}, e_R^{-1}, \begin{pmatrix} \nu_\mu \\ \mu_L \end{pmatrix}^{-1/2}, \mu_R^{-1}. \quad (65)$$

The hypercharges are shown as superscripts.

Because we have two types of gauge bosons, one coupled to Pauli matrices ($SU(2)$) and the other to numbers (one-by-one hermitian matrices are genetaors, and hence their exponentials are one-by-one-unitarity: $U(1)$), this theory is called $SU(2) \times U(1)$. The hypercharge assignments look very bizzarre, but apart from that, everything works fine. We get the $V - A$ interaction, we understand the univerality of the weak interaction, and we correctly reproduced the photon. On the other hand, we now predict a new force, mediated by the $Z$-boson, called "neutral-current weak interaction." We do not know the new parameter, the weak mixing angle yet, and we do not know the masses of the $W$- and $Z$-boson yet. But the combination $G_F/\sqrt{2} = g^2/8m_W^2 = e^2/8m_W^2\sin^2\theta_W$ is already fixed. Once we know the weak mixing angle, we will know the mass of the $W$-boson.