

The Curious Early History of CKM Matrix

-miracles happen!-

Stephen Lars Olsen

Institute for Basic Science, Daejeon South Korea

CAU High Energy Physics Center, Chung-Ang University, Seoul South Korea

**E-mail: solsensnu@gmail.com*

.....

The 1973 Kobayashi Maskawa paper proposed a compelling link between Cabibbo's flavor-mixing scheme and \mathcal{CP} violation but, since it required the existence of six quarks at a time when the physics community was happy with only three, it received zero attention. However, two years after the paper appeared—at which time it had received a grand total of two citations—the charmed quark was discovered and it finally got some notice and acceptance. After this stumbling start, it subsequently emerged as the focal point of an enormous amount of experimental and theoretical research activity. In an invited talk at a KEK symposium to celebrate the 50th anniversary of the KM paper, I reviewed some of the less well known circumstances that occurred in the years preceding and following the paper's appearance.

Some spoilers:

- Kobayashi and Maskawa (and a number of other Japanese physicists) were convinced about the existence of the charmed quark nearly three years before its “discovery” at Brookhaven and SLAC.
- The matrix provided in their seminal 1973 paper was mathematically incorrect. Another version that was in common use for the following twelve years was technically correct, but not really a rotation matrix.
- The CKM matrix \mathcal{CP} phase was only measurable because of the very specific hierarchy of the flavor mixing angles and meson masses.
- Similarly, the neutrino mixing discovery, and the PMNS-matrix measurability were only possible because of favorable values of the neutrino mass differences and mixing angles.

In addition I include some speculations about what may be in store for the future.

.....

Subject Index Quark flavor mixing, Neutrino flavor mixing, \mathcal{CP} violation

arXiv:2309.06042v3 [hep-ph] 11 Jan 2024

1 Introduction

The challenge of reviewing a subject that is fifty years old to a community of experts is to find something to say that isn't already well known to everyone in the audience. However, this obvious truth didn't occur to me when I was invited by the organizers to speak at the KEK special symposium to celebrate the fiftieth anniversary of the Kobayashi-Maskawa six-quark model. An invitation that, in a reckless capitulation to my vanity, I immediately accepted. Upon subsequent reflection, I realized my dilemma: there was precious little that I could say about the hundreds of CKM-related published Belle results—which I expect the organizers had in mind when they offered this invitation—that wasn't already very familiar to the symposium participants. So, instead, I decided to exploit the one advantage I might have over most other participants, and that was that I would be the oldest, or least one of the oldest, person in attendance and reminisce about the early days of the KM era, including some of its pre-history. So, with the forewarning that all historical accounts suffer from mistakes and oversimplifications, and are varnished to match the preconceptions and prejudices of the chronicler, here goes:

2 Prehistory: Cabibbo flavor-mixing and the discovery of \mathcal{CP} violation

The prehistory started sixty years ago during the 1963-64 academic year¹ when there were three major discoveries that all played a major roles in the Kobayashi-Maskawa story: flavor-mixing, quarks, and the observation of \mathcal{CP} violation in $K_L \rightarrow \pi^+ \pi^-$ decays.

2.1 Cabibbo flavor mixing

In their classic paper that identified the $V-A$ coupling of the weak interaction [1], Feynman and Gell-Mann proposed that the weak interaction was a current-current interaction where the hadron current has the form

$$J_\mu = g[\alpha(V_\mu^{\Delta S=0} - A_\mu^{\Delta S=0}) + \beta(V_\mu^{\Delta S=1} - A_\mu^{\Delta S=1})], \quad (1)$$

where g is a coupling constant, $V_\mu^{\Delta S=0}$ and $A_\mu^{\Delta S=0}$ are the vector and axial vector currents for strangeness conserving processes and $V_\mu^{\Delta S=1}$ and $A_\mu^{\Delta S=1}$ are corresponding currents for $\Delta S = \pm 1$ transitions. They also made two additional conjectures. One was *universality*, the notion that

¹ This happened to coincide with my first year as a graduate student at the University of Wisconsin.

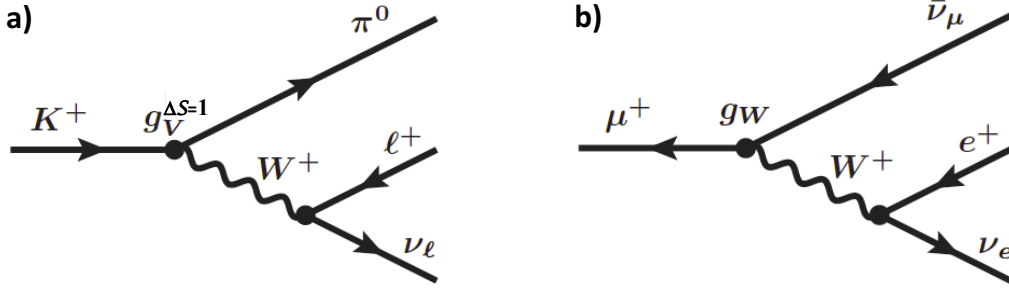


Fig. 1: **a)** (upper) The lowest order Feynman diagram for $K^+ \rightarrow \pi^0 \mu^+ \nu$. **b)** The lowest-order Feynman diagram for $\mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu$.

the currents for the $\Delta S = 0$ and $\Delta S = \pm 1$ hadronic transitions and the

$$g_W(\bar{\nu}_e \gamma_\mu (1 - \gamma_5) e^-) \quad \text{and} \quad g_W(\bar{\nu}_\mu \gamma_\mu (1 - \gamma_5) \mu^-) \quad (2)$$

lepton currents all have a common coupling strength, *i.e.*, $g = g_W$, and $\alpha = \beta = 1$ in eqn. 1, where g_W is related to the square root of the Fermi constant G_F by

$$G_F = \frac{\sqrt{2}}{8} \left(\frac{g_W}{M_W} \right)^2. \quad (3)$$

The other one was the so-called *Conserved Vector Current* (CVC) hypothesis that says that the hadronic matrix elements for the vector component of the weak interaction current are the same as those for the electromagnetic interactions. This has the consequence that vector form-factors for weak decays of hadrons at zero squared momentum-transfers are unity, $f_V(q^2=0) = 1$. These two conjectures translated into a prediction that the coupling strength extracted from the vector-mediated semileptonic process $K^+ \rightarrow \pi^0 e^+ \nu_e$, *i.e.*, $g_V^{\Delta S=1}$ shown in Fig.1a) should be the same as g_W in $\mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu$.

In a paper that appeared in June 1963 [2], Cabibbo pointed out that Feynman-Gell-Mann universality conjecture failed miserably. His comparison of experimental measurements of the partial width for the $\Delta S = 1$ vector weak-interaction process $K^+ \rightarrow \pi^0 \ell^+ \nu$ [3] to the well known width for muon decay found

$$\frac{g_V^{\Delta S=1}}{g_W} \approx 0.26 \quad (4)$$

and about a factor of four below expectations. He also found a similar deviation from universality in the ratio of the axial-vector-mediated partial decay widths $\Gamma(K^+ \rightarrow \mu^+ \nu) / \Gamma(\pi^+ \rightarrow \mu^+ \nu)$:

$$\frac{g_A^{\Delta S=1}}{g_A^{\Delta S=0}} \approx 0.26. \quad (5)$$

(Although the axial-vector currents are not “protected” by CVC, corrections to them were expected to be small [4], and certainly not large enough to account for a factor of four.)

Cabibbo proposed modifying the Feynman-Gell-Mann $\alpha = \beta = 1$ conjecture to $\alpha^2 + \beta^2 = 1$, in which case

$$g_V^{\Delta S=1} = \beta g_W \quad \text{and} \quad \frac{g_A^{\Delta S=1}}{g_A^{\Delta S=0}} = \frac{\beta}{\alpha}, \quad (6)$$

where $\beta \approx 0.25$ could accommodate the abovementioned experimental results. In his paper, Cabibbo proposed his eponymous angle θ_C , which he estimated to $\theta_C \approx 14.9^\circ$, as a convenient way to express two parameters α and β that were subject to the constraint $\alpha^2 + \beta^2 = 1$, and he didn't mention anything about rotations. The earliest experiments that addressed Cabibbo's hypothesis [5] were focused on testing the validity of Cabibbo's relation, $\alpha^2 + \beta^2 = 1$.

The notion that this might represent a rotation didn't become apparent until the 1970 GIM paper [6] that proposed the c -quark as a way to suppress flavor-changing neutral currents. If one accepts the existence of two quark doublets,

$$\begin{pmatrix} u \\ d \end{pmatrix}_s \implies \begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix}, \quad (7)$$

the Cabibbo d - s mixed quark state $d' = d \cos \theta_C + s \sin \theta_C$ is produced by the application of a 2x2 unitary rotation matrix:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos \theta_C & \sin \theta_C \\ -\sin \theta_C & \cos \theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} = \begin{pmatrix} d \cos \theta_C + s \sin \theta_C \\ -d \sin \theta_C + s \cos \theta_C \end{pmatrix}, \quad (8)$$

and has an orthogonal partner, $s' = -d \sin \theta_C + s \cos \theta_C$. In this formulation, it is apparent that Cabibbo's form of weak universality is the same as Feynman-Gell-Mann universality applied to the rotated d' and s' quarks.²

² In addition to suppressing $\Delta S = \pm 1$ weak interaction couplings relative to that for muon decay by a factor of $\sin \theta_C = 0.2245$, Cabibbo's weak universality predicts that $\Delta S = 0$ couplings are suppressed by a factor of $\cos \theta_C = 0.974$. In fact, nuclear physicists had known since 1955 that the half-life for $^{14}\text{O} \rightarrow ^{14}\text{N} \beta^+ \nu$, a vector-mediated $0^+ \rightarrow 0^+$ nuclear β^+ decay transition, was $\sim 3\%$ longer than the value that was predicted using the g_W value determined from muon decay [7, 8]. In 1960, three years before Cabibbo's paper, this discrepancy was noted in the introductory remarks of a *Nuovo Cimento* article on the axial-vector current by Gell-Mann and Levy [4], together with a footnote that suggested that this might be because the unitarity condition might be, in fact, $\alpha^2 + \beta^2 = 1$, and not the $\alpha = \beta = 1$ condition that was conjectured in the Feynman-Gell-Mann $V-A$ paper. The footnote includes an estimate on the mixing that translates into $\theta \approx 14^\circ$, consistent with—and three years before—Cabibbo's estimate for θ_C based on $\Delta S = 1$ transitions. This may explain why K and M, but not C, were awarded the Nobel prize in 2008.

2.2 Gell-Mann Zweig quarks

During this same year Gell-Mann [9] and Zweig [10] proposed the quark model in which hadrons were comprised of fractionally charge fermionic constituents (Zweig called then “aces”). Gell-Mann’s paper was published in January 1964; Zweig’s paper was never published.³ With rotated quarks, the short-distance weak interaction hadronic currents are the same as those for leptons:

$$\begin{aligned} J_\mu^q &= g_W(\bar{u}\gamma_\mu(1 - \gamma^5)d') + g_W(\bar{c}\gamma_\mu(1 - \gamma^5)s') \\ &= g_W \sum_{i,j} (\bar{u}_i\gamma_\mu(1 - \gamma^5)V_{ij}d_j), \end{aligned} \quad (9)$$

where $(u_1, u_2) = (u, c)$ & $(d_1, d_2) = (d, s)$, and V_{ij} is the eqn. 8 quark mixing matrix. The long distance quark-to-hadron processes are described by form factors.

2.3 Discovery of \mathcal{CP} violation

The Christenson, Cronin, Fitch and Turley discovery of the \mathcal{CP} violating decay mode $K_L \rightarrow \pi^+\pi^-$ was reported in the summer of 1964 [11]. This was a relatively low priority experiment that was not aimed at investigating \mathcal{CP} violation but, instead, was designed to investigate some anomalies in coherent $K_2 \rightarrow K_1$ regeneration measurements that had been reported during the previous year [12]. It failed to qualify for a spot in the main experimental hall of the then, almost new, AGS synchrotron that was occupied by spectrometers specialized for total cross section determinations, and π , K , \bar{p} and μ -proton elastic scattering measurements. Instead, the experimental apparatus was located in a relatively inaccessible area inside the AGS magnet ring that the laboratory technical staff referred to as “Inner Mongolia,”⁴ in a neutral particle line that was essentially a hole in the AGS shielding wall that was pointed at a target located in the accelerator’s vacuum chamber, as illustrated in Fig. 2a. The high flux of γ -rays emerging from the target were attenuated by a 3.8 cm-thick lead block followed by a collimator and a bending magnet that swept charged particles out of the beam aperture. A double-arm spectrometer consisting of tracking spark chambers before and after two vertically bending magnets measured the directions and momenta of charged particles that were produced by K_L meson decays that occurred in a 2 m-long decay volume that was a plastic bag filled with atmospheric pressure helium—a low-budget approximation of a vacuum chamber—as shown in Fig. 2b.

³ The story here is that the head of the CERN theory group in 1964, when Zweig was there on a visiting appointment, thought Zweig’s proposed fractionally charged particles was a crackpot idea and refused to provide him with the clerical and drafting support that was needed to prepare a journal-worthy manuscript in the pre-Latex & computer-graphics era. Gell-Mann won the 1969 Nobel physics prize and by 1976, when the head of the theory group became the Director-General of CERN, Zweig was doing biological research and no longer involved in particle physics.

⁴ In the 1960s diplomatic relations between the U.S. and China were non-existent, and mainland China, including Inner Mongolia, was considered by most Americans to be about as accessible as the far side of the Moon.

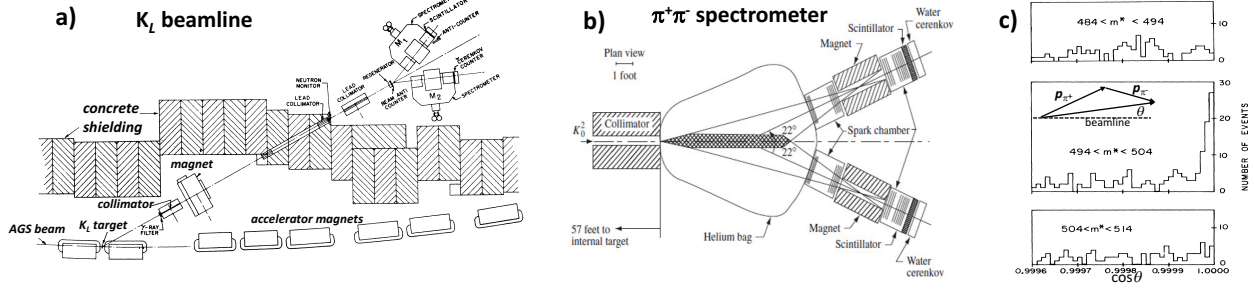


Fig. 2: **a)** The neutral K_L -beamline at the AGS that was used for the $K_L \rightarrow \pi^+ \pi^-$ search experiment. **b)** The two-arm $\pi^+ \pi^-$ spectrometer consisted of a helium-filled decay volume followed by optical tracking spark chambers before and after momentum analyzing magnets **c)** The distribution of events *versus* the cosine of the angle between the direction of the two-track momentum sum and the beamline. The upper, middle and lower panels are for events with two-track invariant masses that are below, centered on, and above m_{K_L} , respectively.

Most of the detected events were due to \mathcal{CP} -allowed $K_L \rightarrow \pi^+ \pi^- \pi^0$ decays and $K_L \rightarrow \pi^\pm \ell^\mp \nu$ ($\ell = e, \mu$) *semileptonic* decays. In these decays, the π^0 or ν was not detected and, as a result, the invariant mass of the two detected charged particles was not, in general, equal to the K_L -meson mass ($m_{K_L} = 498$ MeV). For $K_L \rightarrow \pi^+ \pi^- \pi^0$ decays where the π^0 is undetected, the $\pi^+ \pi^-$ invariant mass is always below 363 MeV; for $K_L \rightarrow \pi^\pm \ell^\mp \nu$, where the ν is missed and the ℓ^\mp track is assigned a pion mass, the two charged track invariant mass distribution ranges from 280 to 546 MeV, with no peak near m_{K_L} . Although the energies of the decaying K_L mesons were not known, their three-momentum directions were confined to be within an rms spread of ± 3.4 mrad ($\pm 0.2^\circ$) around the beamline. A consequence of the missed particle in the three-body decay channels was that the vector sums of the two charge track's momentum vectors did not usually point along the well defined K_L beamline.

Thus, the experimental signature for $K_L \rightarrow \pi^+ \pi^-$ decays was a pair of oppositely charged tracks that, when assigned pion masses, had an invariant mass that was within ± 5 MeV of m_{K_L} and with a summed vector momentum that is directed along the K_L beam direction. Results for these two quantities are shown in Fig. 2c, where the horizontal axis is the cosine of the angle between the $\vec{p}_{\pi^+} + \vec{p}_{\pi^-}$ direction and the K_L beamline, and the upper, central and lower panels show the experimental distributions for $M(\pi^+ \pi^-)$ below, centered on, and above m_{K_L} , respectively. In the central panel there is a pronounced peak totally contained within $\cos \theta > 0.99996$ ($\theta < 9$ mrad), a feature that is absent in the distributions for $M(\pi^+ \pi^-)$ below or above m_{K_L} shown in the upper and lower panels. The ratio of the branching fractions for $K_L \rightarrow \pi^+ \pi^-$ to the sum of all (\mathcal{CP} -conserving) decays to charged particles was $(2.0 \pm 0.4) \times 10^{-3}$.

The signal peak in the central panel of Fig. 2c contained an excess of 45 ± 10 events, but these were not all $K_L \rightarrow \pi^+ \pi^-$ events. About ten of them were due to the coherent $K_L \rightarrow K_S \rightarrow \pi^+ \pi^-$ regeneration process on the helium nuclei in the gas bag decay region, and were indistinguishable

from the $K_L \rightarrow \pi^+ \pi^-$ signal events. Nature was kind. If the branching fraction had been much smaller or the regeneration cross section were higher, the interpretation of the observed signal peak would have been ambiguous. As mentioned above, this was a low-priority experiment. If it had ended up by simply setting an upper limit on the $K_L \rightarrow \pi^+ \pi^-$ branching fraction, who knows when, if ever, a follow-up experiment with higher sensitivity would have occurred.

3 The Kobayashi-Maskawa paper

The famous Kobayashi-Maskawa paper [13] was written in mid-1972, and published in the February 1973 issue of the Japanese journal *Progress of Theoretical Physics*, where it was basically ignored; during the following two and a half years, it received all of two citations. The paper's title is *\mathcal{CP} Violation in the Renormalizable Theory of Weak Interaction*, where the Renormalizable Theory of Weak Interaction is the term they used for what we now call the Standard Model.

Simply put, a \mathcal{CP} violation means that the amplitudes for a processes that involve initial-state particles converting to final-particles and its corresponding antiparticle equivalent are not the same, *e.g.*,

$$\mathcal{M}(a \rightarrow bc) = \langle bc | H_w | a \rangle \neq \bar{\mathcal{M}}(\bar{a} \rightarrow \bar{b}\bar{c}) = \langle \bar{b}\bar{c} | H_w^\dagger | \bar{a} \rangle. \quad (10)$$

But the hermiticity of the Hamiltonian requires that the squares of the amplitudes are equal:

$$|\mathcal{M}(a \rightarrow bc)|^2 = |\langle bc | H_w | a \rangle|^2 = |\bar{\mathcal{M}}(\bar{a} \rightarrow \bar{b}\bar{c})|^2 = |\langle \bar{b}\bar{c} | H_w^\dagger | \bar{a} \rangle|^2, \quad (11)$$

and the only way these two conditions can be satisfied is if \mathcal{M} and $\bar{\mathcal{M}}$ differ by a phase, *i.e.*,

$$\mathcal{M} = |\mathcal{M}| e^{i\delta_{\mathcal{CP}}} \quad \text{and} \quad \bar{\mathcal{M}} = |\mathcal{M}| e^{-i\delta_{\mathcal{CP}}}. \quad (12)$$

So, to incorporate \mathcal{CP} violation into the Standard Model, all you have to do is find a way to insert a complex phase in it somewhere, which, at first glance, wouldn't seem to be so difficult. However this \mathcal{CP} -violating phase is special and, unlike all other phases that show up quantum theories, and including it into the the theory is not at all trivial.

There are countless phases that occur in quantum mechanics; both the Schrödinger and Dirac equations have an imaginary coefficient and their solutions are complex wave functions. But all of these phases, with, so far, only one exception, have the same sign for particles and antiparticles. Only a \mathcal{CP} -violating phase has opposite signs for particles and antiparticles.

The KM paper examines various possible ways that a complex \mathcal{CP} phase might be incorporated into the Standard Model. In the following I discuss the first five pages and the last page separately.

3.1 The KM paper: pages 1→5

In the first five pages, various possibilities were examined and the authors concluded that “no realistic models of \mathcal{CP} -violation exist in the *quartet scheme* without introducing any other new fields.” Here, by the “quartet scheme” they meant the four-quark model that included the charmed quark. Note that this was written in 1971, nearly three years before the c -quark discovery in the “November 1974 revolution.” This, and their page-five conclusion that no realistic model for \mathcal{CP} -violation exists with four quarks raise two questions:

- i) Why were Kobayashi and Maskawa so sure of the existence of the c -quark at such an early date?
- ii) Why can't a \mathcal{CP} -violating phase be introduced together with the Cabibbo angle into the eqn. 8 2×2 quark-flavor mixing matrix?

3.1.1 The discovery of charm: Japanese version

In 1970, a small team of experimenters in Japan led by Kiyoshi Niu, exposed a stack of photographic emulsions to cosmic rays in a high altitude commercial cargo airliner [14]. Upon subsequent inspection they found a remarkable event, shown in Fig. 3, in which an ultra-high energy (multi-TeV) cosmic ray particle interaction produced four charged tracks and two very high energy, closely spaced γ -rays that, when attributed to a $\pi^0 \rightarrow \gamma\gamma$ decay, had a total energy of 3.2 ± 0.4 TeV. Two of the charged tracks, labeled B & C in the figure, have kinks within ~ 5 cm of the production point that are quite distinct in both the X and Y projections shown in Figs. 3a & b, indicating that they decayed to charged daughters (tracks B' & C'). When the event is viewed along the flight direction of track B (Fig. 3c), its daughter charged track (B') and the high energy π^0 are very close to being back-to-back. The transverse momentum of the π^0 relative to the direction of track B was 627 ± 90 MeV, and much higher than was possible for the decay of any known particle at that time. With the π^0 setting the energy scale and assumptions of two-body decays at each kink: $\pi^\pm \pi^0$ for $B \rightarrow B'$ and $\pi^0 p$ for $C \rightarrow C'$ where the secondary π^0 is missed, transverse momentum balance was used to estimate the masses and lifetimes of B and C:

Assumed decay mode	Mass (GeV)	lifetime (sec)
$B \rightarrow \pi^+ \pi^0$	1.78	2.2×10^{-14}
$C \rightarrow (\pi^0) p$	2.95	3.6×10^{-14} ,

The estimated B mass and the proper time intervals are consistent with the GIM estimates of ~ 2 GeV for the charmed-quark mass (and in reasonable agreement with what the now very well determined D^- mass (1.869 GeV) and Λ_c mass (2.286 GeV)). The lifetimes were much shorter than that of any known weakly decaying particle as well as the $\mathcal{O}(10^{-13} \text{ s})$ estimate that was given in the GIM paper. But the latter fact is perhaps not too surprising since emulsion measurements are

biased towards shorter lifetimes. For these reasons, Nagoya theorist Shuzo Ogawa interpreted Niu’s event as being the associated production of an anticharmed meson charmed-baryon pair and their subsequent decays. Although whether or not Niu’s event and Ogawa’s interpretation amounted to a Nobel-prize-worthy claim of a discovery might be a subject of dispute, what matters for our story here is that many people in the Japanese theoretical physics community, especially those in Nagoya that included Kobayashi and Maskawa, were convinced that the charmed quark had been discovered, and that four quarks existed in nature, a scenario that they called the “quartet model.”

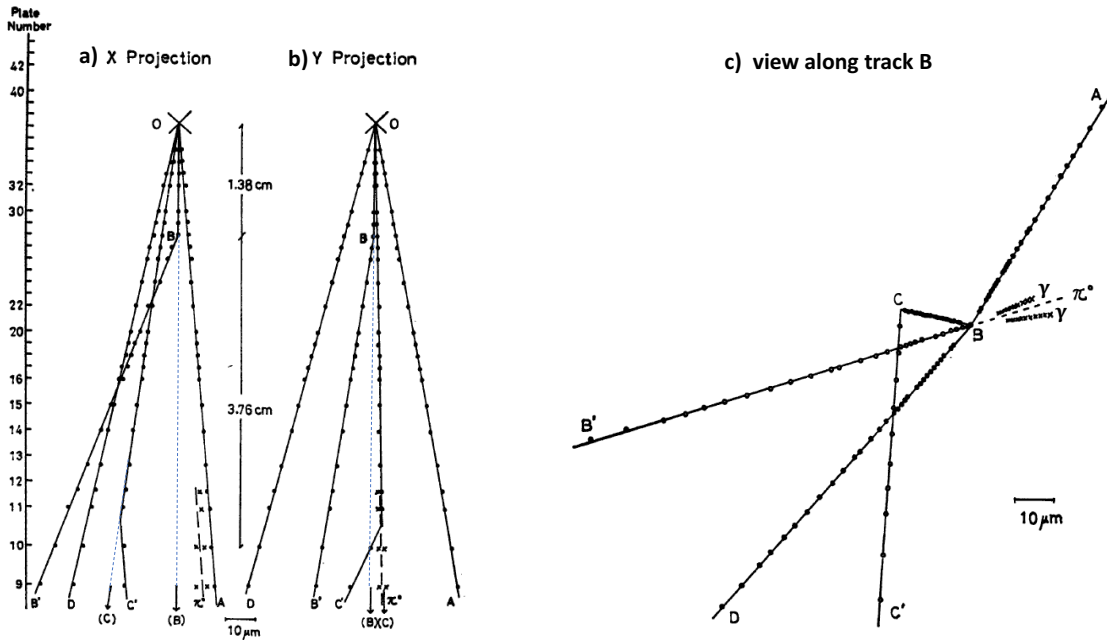


Fig. 3: The a) X and b) Y projections of the Niu event. Here the tracks labeled B and C have kinks at depth of 1.38 cm and 5.14 cm, respectively, that are evident in both views. c) The same event viewed along track B, where the direction of a high energy π^0 , inferred from two detected γ -rays, is very nearly opposite the direction of B’, the daughter track that emerges from the track B kink. (The figures are taken from ref. [14].)

3.1.2 A \mathcal{CP} phase in the four-quark mixing matrix?

In general, a 2×2 matrix like that in eqn. 8 has four complex elements that correspond to eight distinct real numbers. In the four-quark model, flavor mixing is completely described by the single real number θ_C , why can’t one of the other seven numbers be used to specify $\delta_{\mathcal{CP}}$, a \mathcal{CP} -violating phase?

The flavor-mixing matrix describes a rotation and, thus, has to conserve probability. This means it should be unitary: *i.e.* $VV^\dagger = \mathcal{I}$, where \mathcal{I} is the identity matrix:

$$\begin{pmatrix} V_{ud} & V_{us} \\ V_{cd} & V_{cs} \end{pmatrix} \times \begin{pmatrix} V_{ud}^* & V_{cd}^* \\ V_{us}^* & V_{cs}^* \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (13)$$

This corresponds to four relations

$$|V_{ud}|^2 + |V_{us}|^2 = 1 \quad \text{and} \quad |V_{cd}|^2 + |V_{cs}|^2 = 1 \quad (14)$$

$$V_{ud}V_{cd}^* = -V_{us}V_{cs}^* \quad \text{and} \quad V_{cd}V_{ud}^* = -V_{cs}V_{us}^*, \quad (15)$$

that reduces the number of independent parameters from eight to four.

In the weak interaction quark currents (eqn. 9), the total number of quarks is conserved: q_j , which annihilates a q_j -quark, is always accompanied by \bar{q}_i , which creates a q_i -quark. The theory has a subtle property: if each quark field is multiplied by an arbitrary phase factor,

$$d_j \longrightarrow e^{i\phi_j} d_j \quad \text{and} \quad \bar{u}_i \longrightarrow e^{-i\phi_i} \bar{u}_i, \quad (16)$$

and the interactions are modified by the same phases,

$$V \longrightarrow \begin{pmatrix} e^{i\phi_u} & 0 \\ 0 & e^{i\phi_c} \end{pmatrix} \begin{pmatrix} V_{ud} & V_{us} \\ V_{cd} & V_{cs} \end{pmatrix} \begin{pmatrix} e^{-i\phi_d} & 0 \\ 0 & e^{-i\phi_s} \end{pmatrix}, \quad (17)$$

there is no net effect on on the J_μ^q current:

$$\begin{aligned} (\bar{u}_i \gamma_\mu (1 - \gamma_5) V_{ij} d_j) &\longrightarrow (\bar{u}_i e^{-i\phi_i} \gamma_\mu (1 - \gamma_5) e^{i(\phi_i - \phi_j)} V_{ij} e^{i\phi_j} d_j) \\ &= (\bar{u}_i \gamma_\mu (1 - \gamma_5) V_{ij} d_j). \end{aligned} \quad (18)$$

This process is called *rephasing* and the four phases can be expressed as three independent phase differences plus one overall phase that has no effect on the physics. Thus, of the the eight real numbers we started with, four are removed by the unitarity constraint, and three can have any value with no net effect. Thus there is only one number left to define the matrix, and that is needed for the rotation angle θ_C . There is no freedom to add a \mathcal{CP} phase and is what led Kobayashi and Maskawa to conclude that there was no way to incorporate \mathcal{CP} violation into the four quark model.

3.2 The KM paper: page 6

In his 2008 Nobel prize lecture [15], Maskawa recalled that he and Kobayashi had completed the work that was covered in pages 1-5 of their paper and were disappointed with their failure to find any way to incorporate a \mathcal{CP} phase into the four quark model, and were reconciled to the unhappy

likelihood that they would have to publish the negative result. Then, one evening, while—as is customary in Japan—he was taking his after-dinner bath, he mentally went through the calculation described in the previous paragraph, except this time for a six-quark scenario with a 3×3 flavor-mixing rotation matrix in three dimensions. In this case, there are 9 complex elements that are described by 18 real numbers. Of these, 9 are constrained by the unitarity requirement, 5 are taken up by rephasing, and 3 are needed to specify the 3-dimensional rotation,⁵ and that left one number that remained available.

Eureka! With six-quarks there is room for a \mathcal{CP} -violating phase!

A sixth page was added to the manuscript that included the remarks:

Next we consider a 6-plet model, another interesting model of \mathcal{CP} violation,... with a 3×3 instead of 2×2 unitary matrix. In this case we cannot absorb all phases of matrix elements into the phase convention and can take, for example, the following expression:

$$\begin{pmatrix} \cos \theta_1 & -\sin \theta_1 \cos \theta_3 & -\sin \theta_1 \sin \theta_3 \\ \sin \theta_1 \cos \theta_2 & \cos \theta_1 \cos \theta_2 \cos \theta_3 - \sin \theta_2 \sin \theta_3 e^{i\delta} & \cos \theta_1 \cos \theta_2 \sin \theta_3 + \sin \theta_2 \cos \theta_3 e^{i\delta} \\ \sin \theta_1 \sin \theta_2 & \cos \theta_1 \sin \theta_2 \cos \theta_3 + \cos \theta_2 \sin \theta_3 e^{i\delta} & \cos \theta_1 \sin \theta_2 \sin \theta_3 - \cos \theta_2 \sin \theta_3 e^{i\delta} \end{pmatrix}. \quad (19)$$

Then we have \mathcal{CP} -violating effects ... that appear only in the $\Delta S \neq 0$ non-leptonic processes and semi-leptonic decay of neutral strange mesons (we are not concerned with higher states with the new quantum number) ...

(Here, θ_1 is (approximately) the Cabibbo angle, θ_2 is the mixing angle between the 2nd- & 3rd-generation quarks, θ_3 mixes the 1st- & 3rd-generations, and δ is the \mathcal{CP} -violating phase.) And that was it. Discussion about the six-quark model was confined to one paragraph that, together with the expression for the matrix, occupied only about half of the page.

Thus, Kobayashi and Maskawa discovered the way to incorporate a \mathcal{CP} violating phase into the Standard Model, and established a deep theoretical connection between quark-flavor mixing and \mathcal{CP} violation, two subjects that had previously been considered to be unrelated. But this came at a cost: you need to have six-quarks. In 1971, thanks to Kiyoshi Niu, this was not such a big stretch for Kobayashi and Maskawa, who were among the fortunate few who were already convinced that there were (at least) four quarks. On the other hand, most of the world-wide particle physics community outside of Japan was quite satisfied with three quarks.

⁵ In general, for N quark generations with an $N \times N$ mixing matrix, there are N^2 elements characterized by $2N^2$ real numbers. Of these N^2 are used for unitarity, $2N-1$ for rephasing and $N(N-1)/2$ are needed to define the rotation angles. The remaining degrees of freedom that can show up as \mathcal{CP} phases are $(N-1)(N-2)/2$, which vanishes for $N=2$, but is 1 for $N=3$ (and would be 3 if $N=4$).

3.2.1 The first proposal for (and the naming of) charm

The distinction between electron- and muon-neutrinos had been established in 1962 [16], and, in what was at that time the beginnings of the Standard Model, the electron and muon and their neutrinos occupied two weak-isospin=1/2 spinors. In a 1964 paper, Bjorken and Glashow [17] discussed an expansion of $SU(3)$ to $SU(4)$ with the addition of another strangeness-like flavor quantum number. They formulated their model in terms of the Sakata model, the predecessor of the quark model that used the proton-neutron isospin doublet and isoscalar Lambda as basic constituents, with the Lambda replaced by a doublet that had a fourth baryonic constituent with a non-zero value of the new quantum number. When reformulated in the context of the quark model, which had just emerged at that time with three quarks, this was equivalent to adding a fourth quark with matching patterns for the leptons and quarks:

$$\begin{array}{cc} \text{leptons} & \text{quarks} \\ \begin{pmatrix} e^- \\ \nu_e \end{pmatrix} \quad \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix} & \begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix}. \end{array} \quad (20)$$

While their proposal was not very different from schemes that other authors had proposed around that time (see, for example refs. [18–20]), Bjorken and Glashow gave the new flavor the charismatic name “charm,” and that’s the one that stuck; the fourth quark was known as the “charmed quark” (and not the grammatically awkward “charm quark”) even before it was discovered. In retrospect, this four-quark scheme seems like a pretty sound—almost obvious—idea;⁶ how could anyone dismiss such a simple and sensible suggestion? Nevertheless, for the next six years this idea didn’t go very far. By the time the GIM paper [6] appeared in 1970, the Bjorken-Glashow paper had received a grand total of six citations. For some reason, there seem to have been a special affinity in the physics community for the number three and an aversion to the number four.⁷ Moreover, even the GIM paper that proposed a four-quark scenario as a compelling explanation for the suppression of flavor-changing neutral-currents, an important theoretical issue at that time, didn’t experience a boom in citations until after the J/ψ discovery, which finally convinced the world-wide physics community that there were (at least) four quarks.

In 1975, soon after the J/ψ discovery, Maiani (the “M” in GIM), who, at that time, was unaware of the KM result,⁸ wrote an interesting paper [21] that contained all of the KM model and then

⁶ In a passing comment in his original paper on quarks [9], Gell-Mann mentioned a four-quark scheme that was “parallel with the leptons” as an interesting possibility.

⁷ In east Asian cultures the number four is considered to be very unlucky because the Chinese pronunciations of their words for “four” and “death” are homophonous. As far as I know there is no such taboo in Western cultures.

⁸ See footnote 2 in Maiani’s paper.

some. But, thanks to their three-year-long head start, it was Kobayashi and Maskawa—and not Maiani—that got to go to Sweden in December 2008.

The six-quark era started in earnest in 1977, after the discovery of the τ -lepton [22] and the b -quark [23], and the KM model was elevated to role of being the most likely mechanism for \mathcal{CP} violation. Although the b -quark’s charge= $2/3$ partner, the t -quark, wasn’t discovered until 1995 [24, 25], there was very little doubt about its existence.⁹ The only question was its mass; based on the mass pattern of the known quarks, the general consensus was that it was probably in the ~ 25 – 35 GeV range [30].

3.3 “Discovery” of the KM paper

As mentioned above, for over two years the KM paper remained completely unnoticed outside of Japan (and only received limited attention inside Japan). It was finally brought to the attention of the world-wide physics community in a curious set of circumstances that are recounted here.

But first it should be noted that version of the quark-mixing matrix that appeared in the KM paper and is reproduced above in eqn. 19 contains a pretty obvious typographical (or transcription) error. Since the KM matrix nominally describes a rotation, if all three mixing angles and the \mathcal{CP} phase are set to zero, it should revert to the identity matrix, *i.e.*,

$$\mathbf{V} \xrightarrow{(\theta_{1,2,3},\delta)\rightarrow 0} \mathcal{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (21)$$

However in the matrix that appears in the KM paper, the zero-angle limit for the lower-right diagonal element V_{tb} , which should be $V_{tb}\rightarrow 1$, is incorrect:

$$V_{tb}^{\text{KM}} = \cos \theta_1 \sin \theta_2 \sin \theta_3 - \cos \theta_2 \sin \theta_3 e^{i\delta} \xrightarrow{(\theta_{1,2,3},\delta)\rightarrow 0} 0. \quad (22)$$

The first paper to establish that the KM model could account for all that was known about \mathcal{CP} violations at that time, was by Pakvasa and Sugawara and published in the July 1976 issue of Physical Review D [31]. In their paper, they pointed out that in the six-quark model, a non-zero value of ε , the neutral K -meson mass-matrix \mathcal{CP} violation parameter, of the correct magnitude would be produced by the interference between the virtual c - and t -quark contributions to the K^0 - \bar{K}^0 mixing box-diagram shown in Fig. 4a. They also pointed out that in the KM picture, the penguin diagram¹⁰ shown in Fig. 4b, that would mediate direct- \mathcal{CP} violating $K_2\rightarrow\pi^+\pi^-$ decays,

⁹ There were some papers on “topless models,” [26–28] including one with the title: “Does bare-bottom rule out the topless E6 model?” [29], but these did not get much attention.

¹⁰ This was before penguin diagrams were called penguin diagrams.

i.e., the ε' parameter, would be small and consistent with the then existing experimental limits. Since PRD is a widely distributed physics journal, this paper provided the first awareness of the KM paper to the international particle physics community.

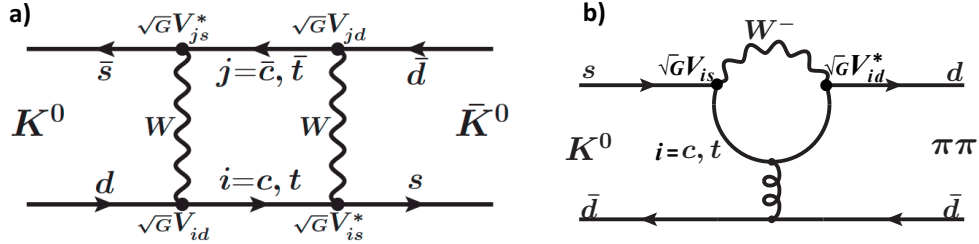


Fig. 4: **a)** The W -exchange Standard Model box diagram for K^0 - \bar{K}^0 mixing (not shown is the one with heavy quark exchange). In the KM six-quark model, the kaon’s mass-matrix \mathcal{CP} -violating parameter ε is produced by interference between the c - and t -quark contributions. **b)** The penguin diagram for direct- \mathcal{CP} -violating $K_2 \rightarrow \pi\pi$ decays.

In their paper, Pakvasa and Sugawara made no mention of the typo in the KM paper and included an expression for the matrix that they attributed to KM but, in fact, was different. This paper also had a typo that mistakenly identified Toshihide Maskawa as K. Maskawa in their citation to the KM paper. Interestingly, many of the papers that immediately followed the Pakvasa-Sugawara paper used the Pakvasa-Sugawara version of the CKM matrix with no mention of the mistake, and with citations to the KM paper that had T. Maskawa incorrectly listed as K. Maskawa (see, it e.g., refs. [32–35]). This recurrence of the typo in the citations provided pretty clear evidence that the Pakvasa-Sugawara PRD article was the source that researchers used for both the matrix and the citation, and that the PTP paper itself was probably not very widely read.¹¹

But, in addition to the typo in their citation to the KM paper, the Pakvasa-Sugawara version of the KM matrix had a problem of its own. In their paper, the KM expression for V_{tb} , given above in eqn. 22, was replaced by

$$V_{tb}^{\text{PS}} = \cos \theta_1 \sin \theta_2 \sin \theta_3 - \cos \theta_2 \cos \theta_3 e^{i\delta}, \quad (23)$$

which doesn’t go to zero in the limit of zero mixing angles. But neither does it go to 1, instead,

$$V_{tb}^{\text{PS}} \xrightarrow{(\theta_{1,2,3}, \delta) \rightarrow 0} -1, \quad (24)$$

¹¹ But why didn’t Pakvasa and Sugawara alert their readers about the problem with the matrix in the KM paper? Sugawara’s recollection is that when he first learned about the KM paper at a University of Tokyo physics seminar, he was impressed by their six-quark scheme and “reconstructed it in [his] own way, without reading their paper carefully.” When he and Pakvasa subsequently did their analysis and wrote up their results, they included Sugawara’s version of the matrix, which was the KM version without the error, in their paper. In fact, Pakvasa and Sugawara remained unaware of the KM paper’s typo. According to Sugawara, “I never realized that the paper had this typo until it was recently pointed out to me.”

and for $\delta = 0$, has a negative determinant. Nevertheless, this version of the matrix is unitary, which is the only essential requirement for a quark-flavor mixing matrix, and this form was used in much (but not all) of the literature until 1984, when the currently widely accepted Chau-Keung parameterization was proposed and soon thereafter endorsed by the Particle Data Group.

4 Reparameterizing the CKM matrix

A rotation in three dimensions can be accomplished by three successive rotations: first by an angle θ_1 around the z axis that mixes x and y $[(x, y) \rightarrow (x', y')]$, then by θ_2 about the new x' axis that mixes y' and z $[(y', z) \rightarrow (y'', z')]$ and, finally, by θ_3 around the y'' axis that mixes x' and x'' $[(x', z') \rightarrow (x'', z'')]$. This is just one of many ways that can be used to specify a given rotation. For example, the order of the three rotations can be changed, and there is freedom in the choice of the axes that are used to define the rotations. For these different choices, the values of the individual rotation angles are different, as are the expressions for each matrix element in terms of these rotation angles. Ultimately, however, the numerical value of the magnitude of each matrix element for any of these choices has to be the same, and independent of the choice of the individual rotations. In addition, in the case of the CKM matrix, which is complex, rephasing invariance provides five independent arbitrary phase parameters that can be attached to the various matrix elements to establish whatever phase convention may seem convenient. The physics content is independent of these parameterization choices.

On what basis should a parameterization be selected? In answer to this, Haim Harari suggested some criteria for what he would consider to be a “good” parameterization. These included [36]:

- There should be a simple relation between the most directly measurable matrix elements V_{ij} and the quark mixing angles.
- The matrix elements above the diagonal, which correspond to kinematically allowed decay processes that are directly measurable, should have the simplest possible expressions.
- If possible, the \mathcal{CP} violating phase should be linked to only one angle, and preferably the sine of that angle.

During the years immediately following the wide recognition of the KM paper, there was considerable effort aimed at finding a suitable parameterization. This was aided by concurrent experimental measurements of the relative magnitudes of the V_{cb} and V_{ub} matrix elements using B -mesons that were produced via e^+e^- annihilations at two colliders that existed at that time: PEP at SLAC and CESR at Cornell.

4.1 Experimental information about quark transitions

The PEP project was initially conceived as a two ring proton-electron-positron collider with an electron-positron ring that could support $E_{\text{cm}}=30$ GeV e^+e^- collisions primarily to search for the top-quark (if its mass was less than 15 GeV), and a second 150 GeV proton ring that could support $e^\pm p$ collisions with $E_{\text{cm}} \approx 100$ GeV for measurements of deep inelastic scattering at high energies and Q^2 values. The CESR collider was conceived during 1974 as a follow-up to the Cornell fixed-target e^-p scattering program, and proposed to the U.S. National Science Foundation in 1975, soon after the J/ψ discovery, as an $E_{\text{cm}} \approx 16$ GeV e^+e^- collider primarily aimed at studies of charmed particles. The PEP and CESR projects were both well underway when the b -quark was discovered in 1977. Fortuitously, the initial CESR design energy could comfortably operate in the $E_{\text{cm}}=9-11$ GeV range, and cover the three narrow $\Upsilon(1S, 2S, 3S)$ resonances and the threshold region for $e^+e^- \rightarrow B\bar{B}$ meson pair production, that was expected to be around $E_{\text{cm}}=10.5$ GeV.¹²

The two projects started running in 1979 and soon thereafter provided convincing experimental evidence for a strong hierarchy among the weak interaction mixing angles for the three quark generations. It was already well established that transitions *within* the first quark generation, *e.g.*, $u \rightarrow d$ and *within* the second generation ($c \rightarrow s$) were strongly favored over transitions *between* the first and second generations ($s \rightarrow u$ and $c \rightarrow d$), *i.e.*, the Cabibbo angle. Experiments at PEP found that transitions *between* the second and third generation ($c \rightarrow b$) are more suppressed those between the first and second generations, and the CESR experiment determined that transitions between the first and third generations ($u \rightarrow b$) are the least favored of all.

1st \leftrightarrow 2nd generation: The suppression of strangeness-changing decays that was noted by Cabibbo in 1963, led to the realization that the relative strengths of the $u \rightarrow d$ and the strangeness-changing $u \rightarrow s$ transitions are modulated by factors of $\cos \theta_C$ and $\sin \theta_C$, respectively, where $\theta_C = 13^\circ$ is the Cabibbo angle. Thus the diagonal element $V_{ud}=0.974$ is nearly unity and almost five times larger than its adjacent entry $V_{us}=0.225$.

2nd \leftrightarrow 3rd generation: In 1983, the MAC experiment at the PEP collider used b -flavored hadrons (mainly B^\pm and B^0 mesons) produced via the $e^+e^- \rightarrow b\bar{b}$ annihilation process (about 1/10th of the total annihilation cross section) to determine the b -quark lifetime by measuring the impact parameters of charged leptons from $B \rightarrow X\mu\nu$ and $Xe\nu$ inclusive semileptonic decays, where X is a hadronic system [37]. (The definition of the impact parameter is indicated in Fig. 5a.) This was a difficult measurement: the parent B -meson's energy and direction were not precisely known on an event-by-event basis; there was contamination from semileptonic decays

¹²The CESR facility spent its life operating at $E_{\text{cm}} \sim 10-11$ GeV.

of charmed mesons; and the mean value of the impact parameter that was eventually determined ($\approx 170 \mu\text{m}$) was substantially smaller than the experimental resolution ($\sim 600 \mu\text{m}$) as well as the horizontal size of the beam-beam interaction region ($\sim 400 \mu\text{m}$). Nonetheless, the measured impact parameter distributions for muons and electrons shown in the upper and lower panels of Fig. 5b, respectively, both had excesses at positive values and these translated into a b -quark lifetime of $\tau_b = 1.8 \pm 0.7$ ps. Since this was about a factor of four times longer than the lifetime of the much lighter c -quark, it was a big surprise at that time. As discussed below, $b \rightarrow c$ transitions are the b -quark's dominant decay mechanism, and the measured lifetime could be used to determine a value for $|V_{cb}|$ using the expression

$$\Gamma_b = 1/\tau_b = |V_{cb}|^2 \frac{G_F^2 m_b^5 c^4}{192\pi^3 \hbar^7} (2_{\text{quarks}} \times 3_{\text{colors}} + 3_{\text{leptons}}), \quad (25)$$

which is the (textbook) expression for the muon lifetime tailored to the b -quark mass, modified to include the effect of $|V_{cb}|$ on the coupling strength, and multiplied by the number of accessible final states: two quark flavors, three quark colors, and three types of leptons, as illustrated in Fig. 5c. The result was $|V_{cb}| \approx 0.04$, and about a factor of five smaller than $|V_{us}|$, a difference that is similar to (but not exactly the same as) the factor of five Cabibbo suppression¹³ between V_{us} and V_{ud} . The MAC results were confirmed by the MarkII and DELCO experiments at PEP [39, 40] and the TASSO experiment [41] at PETRA, an $E_{\text{cm}}=40$ GeV e^+e^- collider at the DESY laboratory in Hamburg.

$1^{\text{st}} \leftrightarrow 3^{\text{rd}}$ generation: The CLEO experiment studied semileptonic $B \rightarrow X \ell \nu$ decays of B mesons that were produced at $E_{\text{cm}} = 10.58$ GeV, the peak¹⁴ of the $\Upsilon(4S) \rightarrow B\bar{B}$ resonance [42]. Since this energy is only 20 MeV above the $B\bar{B}$ threshold, the B mesons are produced very nearly at rest (the boost factor is $\gamma\beta = 0.062$) and the energy of a decay lepton ($\ell = e, \mu$) in the laboratory frame is very nearly equal to what it is in the B meson rest frame. In $b \rightarrow c \ell \nu$ -mediated decays, the minimum mass of the hadronic system is $M_X^{\text{min}} = m_D = 1.86$ GeV, and this translates into a maximum lepton momentum of $p_{\ell\nu}^{\text{max}} = 2.31$ GeV/ c ; in $b \rightarrow u \ell \nu$ -mediated decays, the mass of the hadronic system can be as light as $M_X^{\text{min}} = m_\pi$, and the end-point momentum is $p_{u\ell\nu}^{\text{max}} = 2.60$ GeV/ c . Thus, measurements of the lepton momentum spectra in the end-point region can be used to determine the relative strengths of the $b \rightarrow c$ and $b \rightarrow u$ transitions and extract the value of $|V_{ub}|^2/|V_{bc}|^2$. Figure 6 shows the measured momentum distributions for electrons (*upper*) and muons (*lower*) together with expectations for $b \rightarrow c \ell \nu$ (dashed curves)

¹³ The PDG 2020 [38] value for the B meson lifetime is $\tau_b = 1.519 \pm 0.004$ ps and that for the V_{cb} matrix element is $|V_{bc}| = 0.0410 \pm 0.0014$, which are both within the error ranges of the MAC measurements.

¹⁴ The $\Upsilon(4S)$ is the 3rd radial excitation of the $\Upsilon(1S)$, the lowest-lying $J^{PC} = 1^{--}$ “bottomonium” ($b\bar{b}$) meson, and the first that is above the $B\bar{B}$ “open bottom” threshold. At its peak, the cross section for $e^+e^- \rightarrow B\bar{B}$ is about 1 nb.

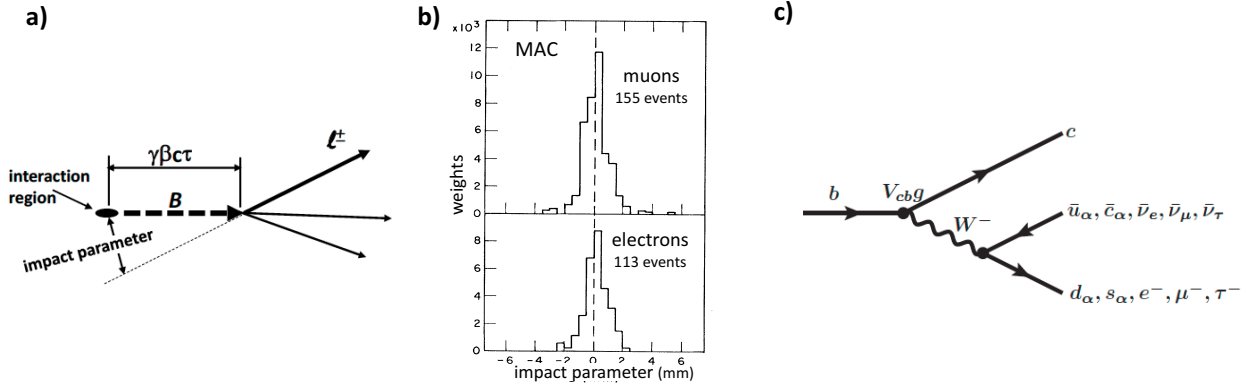


Fig. 5: a) A sketch of the projections of tracks from the semileptonic decay of a B meson onto a plane perpendicular to the e^+e^- beamline, together with indications of the size of the beam-beam interaction region, and the definition of the impact parameter, *i.e.*, the distance of closest approach of the charged lepton track to the e^+e^- interaction point. The unknown B meson's direction was assigned with reasonably good accuracy to be along the event's thrust axis. b) Impact parameter distributions for muons (upper) and electrons (lower) (from the MAC experiment [37]). Here each entry is weighted by the inverse square of its measurement error. c) A quark-line diagram for b -quark decays. The subscript α indicates the quark color.

and $b \rightarrow ul\nu$ (dotted curves). There are no events in the $2.31 < p_{\text{lepton}} < 2.60$ GeV/ c range that could be unambiguously attributed to $b \rightarrow ul\nu$ decays and the shapes of the spectra are consistent with expectations for $\sim 100\%$ $b \rightarrow cl\nu$ with no significant contribution from $b \rightarrow ul\nu$. From these data, the CLEO group established a 90% CL upper limit¹⁵ of $|V_{ub}|/|V_{cb}| < 0.14$.

4.2 The Chau-Keung and Wolfenstein parameterizations

Since the early parameterizations fail on all three of the Harari criteria, a better one was needed. Of a number of proposed replacements[21, 36, 43–45], two continue to be widely used today: one by Chau and Keung [43] and the other by Wolfenstein [44].

4.2.1 Chau-Keung parameterization

The parameterization proposed by Chau and Keung was specifically motivated by the occurrence of the \mathcal{CP} violating phase in the $V_{tb} = \cos \theta_1 \cos \theta_2 \cos \theta_3 - \cos \theta_2 \cos \theta_3 e^{i\delta}$ term in the Pakvasa-Sugawara version of the original KM parameterization that seemed to suggest that there is a large \mathcal{CP} violation that is confined to the (t, b) quark sector. This was in sharp contrast to the results of detailed calculations of measurable \mathcal{CP} violating effects that invariably resulted in small numbers that involved factors of $\sin \theta_2 \sin \theta_3 e^{i\delta}$.

¹⁵ The PDG 2020 [38] value is $|V_{ub}|/|V_{cb}| = 0.093 \pm 0.004$.

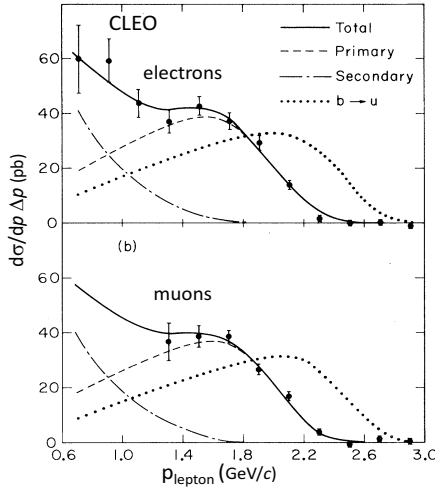


Fig. 6: Lepton momentum spectra for $B \rightarrow X e \nu$ (upper) and $B \rightarrow X \mu \nu$ (lower) decays. The dashed (dotted) curves are expectations for $b \rightarrow c l \nu$ ($b \rightarrow u l \nu$). The dash-dot curves shows the expectation for secondary leptons from semileptonic decays of D decays produced via $B \rightarrow D l \nu$ and of τ leptons from $B \rightarrow X \tau \nu$. (From the CLEO experiment [42].)

q

(For clarity, in the following we follow the more transparent PDG notation that uses θ_{ij} to denote the mixing angle around an axis that is perpendicular to the (i, j) plane. The translation between $\theta_{1,2,3}$ and θ_{ij} is: $\theta_1 = \theta_{12}$, $\theta_2 = \theta_{23}$, and $\theta_3 = \theta_{13}$. In addition we abbreviate $\cos \theta_{ij}$ by c_{ij} and $\sin \theta_{ij}$ by s_{ij} .)

The Chau-Keung parameterization, which is exactly unitary, is given by the (right-to-left) sequence of two-dimensional rotations:

$$\begin{aligned}
 V_{\text{CKM}} &= \overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix}}^{\theta_{23} \text{ about } d''} \overbrace{\begin{pmatrix} c_{13} & 0 & s_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13} e^{i\delta} & 0 & c_{13} \end{pmatrix}}^{\theta_{13} \text{ about } s' \text{ (incl. } \delta)} \overbrace{\begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{\theta_{12} \text{ about } b} \\
 &= \begin{pmatrix} c_{12} c_{13} & s_{12} c_{13} & s_{13} e^{-i\delta} \\ -s_{12} c_{23} - c_{12} s_{23} s_{13} e^{i\delta} & c_{12} c_{23} - s_{12} s_{23} s_{13} e^{i\delta} & s_{23} c_{13} \\ s_{12} s_{23} - c_{12} c_{23} s_{13} e^{i\delta} & -c_{12} s_{23} - s_{12} c_{23} s_{13} e^{i\delta} & c_{23} c_{13} \end{pmatrix}, \tag{26}
 \end{aligned}$$

where it is apparent that this version passes the requirement that $V \rightarrow \mathcal{I}$ when $\theta_{ij} \rightarrow 0$. Here Harari's three criteria are satisfied: the $\theta_{12}, \theta_{23}, \theta_{13}$ mixing angles correspond to experimentally distinct $u \leftrightarrow s, c \leftrightarrow b, \& u \leftrightarrow b$ transitions, respectively; the \mathcal{CP} phase factor is always multiplied by a factor containing s_{13} ($= |V_{ub}|$); and the three above-diagonal terms have simple forms. Moreover, in this parameterization, θ_{12}, θ_{23} and θ_{13} are all in the first quadrant, which means that s_{ij} and c_{ij} are all positive, the \mathcal{CP} phase δ is also positive and in the range $0 \leq \delta < 2\pi$, and the matrix reduces to the

2×2 Cabibbo matrix for $\theta_{23} = \theta_{13} = 0$. This has been the PDG parameterization of choice since 1986 [46] and, according to recent measurements [38],

$$\begin{aligned}\theta_{12} &= 13.09^\circ \pm 0.03^\circ & \theta_{23} &= 2.32^\circ \pm 0.04^\circ \\ \theta_{13} &= 0.207^\circ \pm 0.007^\circ & \delta &= 68.53^\circ \pm 0.51^\circ.\end{aligned}\quad (27)$$

4.2.2 Wolfenstein parameterization

The Wolfenstein parameterization is an approximation that employs a polynomial expansion in terms of $\lambda \equiv \sin \theta_C = 0.2265$ that reflects the hierarchical character of the CKM matrix. With an accuracy up to $\mathcal{O}(\lambda^3)$ it has the form:

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \quad (28)$$

where the parameter $A \approx 0.8$ accounts for the fact that the Cabibbo-like suppression between V_{cb} and V_{cd} is about 20% more severe than that between V_{cd} and V_{ud} . In this parameterization, all of the \mathcal{CP} violation resides in the single parameter η that is confined to the V_{cd} and V_{ub} corners of the matrix where it is multiplied by $\lambda^3 A$, a small number.

This parameterization is very convenient and is widely used, but in a somewhat modified form that was suggested by Buras and colleagues [47]. In the Buras version, the Wolfenstein λ , A , ρ and η parameters are redefined in terms of the Chao-Keung angles to be exactly

$$\lambda \equiv s_{12} = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}} \quad (29)$$

$$A\lambda^2 \equiv s_{23} = \lambda \frac{|V_{cb}|}{|V_{us}|} \quad (30)$$

$$A\lambda^3(\rho - i\eta) \equiv s_{13}e^{i\delta} = V_{ub}^*. \quad (31)$$

With these parameter definitions, V_{ub} is the same as in the Chau and Keung parameterizations, and the higher corrections to V_{us} and V_{cb} start at $\mathcal{O}(\lambda^7)$ and $\mathcal{O}(\lambda^8)$, respectively. In addition Buras *et al.* defined new parameters $\bar{\rho}$ and $\bar{\eta}$ as

$$\bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}, \quad (32)$$

in which case

$$\bar{\rho} = \rho\left(1 - \frac{\lambda^2}{2}\right) + \mathcal{O}(\lambda^4) \quad \text{and} \quad \bar{\eta} = \eta\left(1 - \frac{\lambda^2}{2}\right) + \mathcal{O}(\lambda^4) \quad (33)$$

and

$$V_{td} = A\lambda^3(1 - \bar{\rho} + i\bar{\eta}) + \mathcal{O}(\lambda^7). \quad (34)$$

Although the distinction between Wolfenstein’s ρ , η and (the nearly equal) $\bar{\rho}$, $\bar{\eta}$ parameters may seem to be confusing and unnecessary, the latter parameters are preferred and are more generally used (for reasons that are discussed in detail in ref. [48]). The PDG review [38] only provides values for the four (redefined) “Wolfenstein parameters” that, in 2020, were

$$\begin{aligned} \lambda &= 0.22650 \pm 0.00048 & A &= 0.790_{-0.012}^{+0.017} \\ \bar{\rho} &= 0.141_{-0.017}^{+0.016} & \bar{\eta} &= 0.357 \pm 0.011. \end{aligned} \quad (35)$$

The eqn. 28 form of the matrix is carefully tuned for \mathcal{CP} violations in the b -quark sector that is produced by the imaginary parts of V_{ub} and V_{td} , and, since all the matrix elements in the second row and column, *i.e.*, the ones that involve the strange and charmed quarks, are real, it is not applicable to descriptions of \mathcal{CP} violation the c - and s -quark sectors. For this, Wolfenstein provided a version of the matrix that is expanded to include \mathcal{CP} -violating terms of $\mathcal{O}(\lambda^4)$ in V_{ts} and $\mathcal{O}(\lambda^5)$ in V_{cd} :

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\bar{\rho} - i\bar{\eta}) \\ -\lambda(1 + iA^2\lambda^4\bar{\eta}) & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \bar{\rho} - i\bar{\eta}) & -A\lambda^2(1 + i\lambda^2\bar{\eta}) & 1 \end{pmatrix} + \mathcal{O}(\lambda^6). \quad (36)$$

An expression involving the same parameters that is exactly unitary is given by Kobayashi in ref. [49]

5 \mathcal{CP} violation in b -quark decays?

In the KM model, \mathcal{CP} violation in neutral K -meson decays, other than the ε mass-matrix parameter, are mainly produced by complex phases in the upper-left 2×2 corner of the KM matrix¹⁶ that, in Wolfenstein’s $\mathcal{O}(\lambda^5)$ parameterization, (eqn. 36) is the confined to the phase of V_{cd} , and is tiny:

$$\arg(V_{cd}) = A^2\lambda^4\bar{\eta} = 5.9 \times 10^{-4} \approx 0.03^\circ. \quad (37)$$

In contrast, the CKM matrix element for charmless decays of B -mesons that proceed via $b \rightarrow u$ transitions is V_{ub} , with a \mathcal{CP} phase δ that (we now know) is $\delta \approx 70^\circ$. If the kaon’s direct- \mathcal{CP} -parameter ε' , caused by a tiny, fraction of a degree phase, can be measured in the 20th century, the observation of \mathcal{CP} violations produced by a 70° phase in B -meson decay in the 21st century should be easy.

Wrong! There are a number of important differences between the neutral kaon and B -meson systems that make the types of measurements that were used to discover and elucidate the properties of \mathcal{CP} violations in the kaon system inapplicable to the B -meson system.

¹⁶ V_{ub} can contribute to kaon decays via penguin diagrams, but these contributions are suppressed by similar $\mathcal{O}(A\lambda^4)$ factors.

B-mesons have a huge number of different decay channels.

In contrast to the K -meson system, where 99.98% of K_S decays are to either $\pi^+\pi^-$ or $\pi^0\pi^0$ final states, and 99.7% of K_L decays are to either $\pi\ell\nu$ or $\pi\pi\pi$ final states, B -mesons have hundreds of different decay modes almost all of which have, at best, fraction of a percent level branching ratios.

The B^0 - \bar{B}^0 mass eigenstates have very short, and nearly equal lifetimes. In the K -meson system, the K_S and K_L mass eigenstates have large lifetime differences (0.1 ns vs. 52 ns, respectively), and an essentially pure K_L beam can be achieved by simply making a neutral beam line that is longer than several K_S proper decay lengths. In comparison, the equivalent B_H and B_L mass eigenstates have very nearly equal lifetimes of a mere 1.5 ps ($c\tau_B = 0.45$ mm), and there is no possibility for making a beamline of \mathcal{CP} -tagged B -mesons, much less one that distinguishes between the two different \mathcal{CP} values.

B-meson decays to final states that are eigenstates of \mathcal{CP} are infrequent.

K_S mesons decay almost exclusively to \mathcal{CP} -even $\pi^+\pi^-$ and $\pi^0\pi^0$ eigenstates; K_L decays to \mathcal{CP} -odd $\pi^+\pi^-\pi^0$ and $\pi^0\pi^0\pi^0$ eigenstates occur with branching fractions of 19.5% and 12.5%, respectively. In contrast, the most prominent \mathcal{CP} -eigenstate decay mode for neutral B -mesons is $B^0 \rightarrow J/\psi K_S$, with a meager 0.045% branching fraction.

Moreover, as noted above, V_{cb} , which has no \mathcal{CP} -violating phase, has a magnitude that is an order of magnitude larger than V_{ub} and, thus, branching fractions for $b \rightarrow u$ mediated “non-charmed” decays of B mesons are strongly suppressed. As a result the prospects for finding and studying \mathcal{CP} violations in the B -meson system looked pretty hopeless.

5.1 Prospects for testing the KM \mathcal{CP} mechanism: pre 1980

Sometime around 1979-80, Abraham Pais who, along with Murray Gell-Mann was responsible for many of the fundamental theoretical discoveries in the early days of flavor physics, discussed the prospects for \mathcal{CP} measurements with charmed and beauty mesons in a seminar at Rockefeller University in New York City, where he had this to say about \mathcal{CP} violations with heavy quarks [50]:

“There is good news and bad news. The good news is that \mathcal{CP} violation in a heavy meson system is quite similar to that of the K -meson system. The bad news is that there is little distinction like the K_S - K_L mass eigenstates. For heavy meson systems, both lifetimes are short.”

In the audience was a young theorist Ichiro (Tony) Sanda, who recalls thinking at that time [50]:

“‘ \mathcal{CP} violation in a heavy meson system is quite similar to that of the K meson system?’—How could anything as interesting as \mathcal{CP} violation be so uninteresting.”

and he resolved to find a way to prove that Pais was wrong.

5.2 Tony Sanda's great idea

At this same time, I was one of the founding members of the CLEO experiment that was located on the Cornell University campus in upstate New York, about a two-hour drive from New York City.¹⁷ The CESR e^+e^- collider was in its infancy and had a maximum instantaneous luminosity of $\mathcal{L} \sim 5 \times 10^{30} \text{cm}^{-2}\text{s}^{-1}$. We had just discovered the $\Upsilon(4S)$ resonance [51] and while running at $E_{\text{cm}} = 10.58 \text{ GeV}$, its peak energy, we could collect about 30 $\Upsilon(4S) \rightarrow B\bar{B}$ events/day (see Fig. 7a). This was, at that time, the most prolific source of B mesons in the world and we were anxious to make good use of them. To this end, my Cornell colleagues invited Sanda for some seminars. During his first seminar, the best strategy that Sanda had to offer was a vague plan to search for a $\ell^+\ell^+$ vs. $\ell^-\ell^-$ asymmetry in events of the type

$$e^+e^- \rightarrow B^0\bar{B}^0 \rightarrow \ell^\pm\ell^\pm + \text{anything}, \quad (38)$$

with the faint hope that somehow a measurable \mathcal{CP} violation would show up. But this was very much like the frequently performed $K^0(\tau) \rightarrow \pi^-\ell^+\nu$ vs. $K^0(\tau) \rightarrow \pi^+\ell^-\bar{\nu}$ asymmetry measurements, but without any of the above-listed advantages that make the neutral kaon system so special. The experimenters in the audience, who were all hyped up to do great and wondrous things with the B mesons that they had worked so hard to produce, were noticeably disappointed. When Sanda got back to New York City, he felt under strong pressure to come up with something that was new and unique to B mesons.

A few months later, in his second seminar at Cornell, he did just that. He proposed a scheme that he developed in collaboration with Ashton Carter [52, 53] for using interference between the $B^0 \rightarrow K_S J/\psi$ & $B^0 \rightarrow \bar{B}^0 \rightarrow K_S J/\psi$ decay amplitudes that eventually became the primary motivation for the BaBar and Belle asymmetric B -factory experiments and was the basis for Kobayashi and Maskawa's Nobel prize.

The idea, which is illustrated in Fig. 7b, was very elegant. You start with a flavor-tagged B^0 (or \bar{B}^0 —here I use a tagged B^0 to illustrate the idea) that can directly decay via the $B^0 \rightarrow K^0 J/\psi$ diagram shown in the top panel, or decay via the indirect route where it first mixes into a \bar{B}^0 that then decays via $\bar{B}^0 \rightarrow \bar{K}^0 J/\psi$. But experiments don't distinguish between K^0 or \bar{K}^0 decays, instead they measure K_S and K_L decays. Thus, in events where a K_S is detected, the direct and indirect decay routes access identical final states and interfere. (Likewise for events where a K_L is detected, except here the interference term has an opposite sign.)

The direct amplitude has no \mathcal{CP} phase (at least not at leading order), but the indirect amplitude has an extra factor of V_{td}^2 (not $|V_{td}|^2$!) and so, a \mathcal{CP} phase of -2β . For tagged \bar{B}^0 decays,

¹⁷ At that time I was at the University of Rochester, a two-hour drive in the opposite direction.

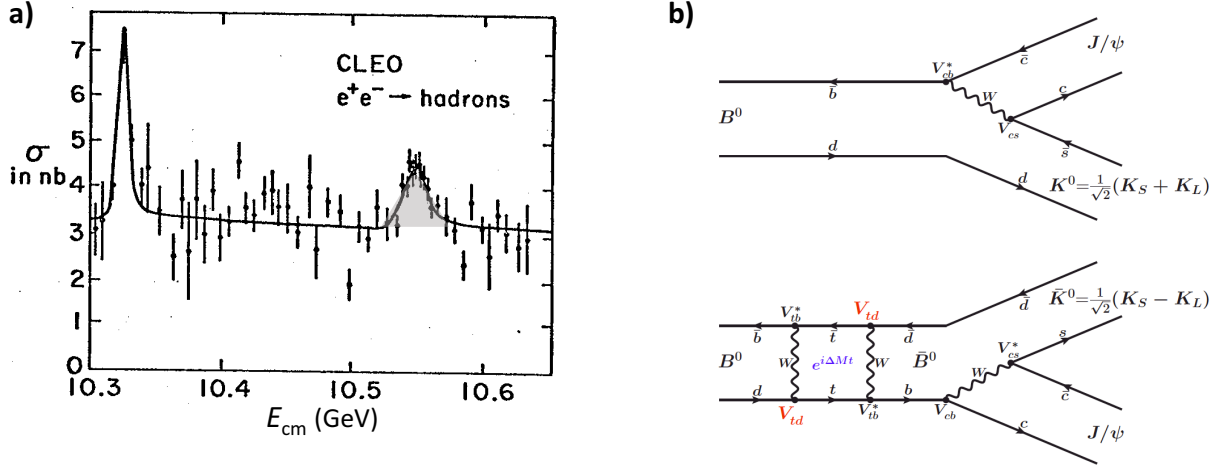


Fig. 7: a) CLEO's 1980 measurements of $\sigma(e^+e^- \rightarrow \text{hadrons})$ in the $\Upsilon(3S)$ & $\Upsilon(4S)$ resonance region (from ref. [51]). b) Tree diagrams for $B^0 \rightarrow J/\psi K^0$; $K^0 \rightarrow K_S$ (or K_L) (upper) and for $\bar{B}^0 \rightarrow J/\psi \bar{K}^0$; $\bar{K}^0 \rightarrow K_S$ (or K_L) after it is produced via $B^0 \rightarrow \bar{B}^0$ mixing (lower). The \bar{B}^0 picks up a common phase $\delta_{\text{com}} = \Delta M \tau$ from the mixing and a \mathcal{CP} violating phase $\phi_{\mathcal{CP}} = -2\beta$ from the two V_{td} vertices in the mixing box.

the CKM factor is V_{td}^{*2} and the \mathcal{CP} phase is $+2\beta$. Thus, the $\bar{B}^0(\tau) \rightarrow f_{\mathcal{CP}}$ vs. $B^0(\tau) \rightarrow f_{\mathcal{CP}}$ time-dependent asymmetry, where $f_{\mathcal{CP}}$ is any \mathcal{CP} -eigenstate is

$$\mathcal{A}_{B \rightarrow f_{\mathcal{CP}}}^{\mathcal{CP}}(\tau) = \frac{\bar{\Gamma}_{\bar{B}^0 \rightarrow f_{\mathcal{CP}}}(\tau) - \Gamma_{B^0 \rightarrow f_{\mathcal{CP}}}(\tau)}{\bar{\Gamma}_{\bar{B}^0 \rightarrow f_{\mathcal{CP}}}(\tau) + \Gamma_{B^0 \rightarrow f_{\mathcal{CP}}}(\tau)} = -\xi_{\mathcal{CP}} \sin(2\beta) \sin(\Delta M_B \tau), \quad (39)$$

where $\xi_{\mathcal{CP}} (= -1$ for $K_S J/\psi$ and $+1$ for $K_L J/\psi)$ is the \mathcal{CP} eigenvalue of $f_{\mathcal{CP}}$, $\Delta M_B = M_H - M_L$ is the mass difference between the neutral B_H and B_L mass eigenstates (*i.e.*, the B^0 - \bar{B}^0 mixing frequency), and τ is the proper time between the $B^0 \rightarrow K J/\psi$ ($B_{\mathcal{CP}}$) decay and the flavor-specific decay of the accompanying \bar{B}^0 meson (B_{tag}), whose decay products are used to tag the B^0 meson's flavor. Note that in e^+e^- colliders, τ can be positive (when the $B_{\mathcal{CP}}$ decay occurs after the B_{tag} decay), or negative (when the $B_{\mathcal{CP}}$ decay occurs first), and the time-integrated asymmetry is zero.

5.2.1 Great idea! but is it practical?

The idea was new, and the mechanism was unique to B mesons, but there were many pieces that had to fall into just the right places for Sanda's proposal to have any chance of being practical. Since at that time there was no experimental information about the b -quark-related CKM elements or B^0 - \bar{B}^0 mixing, there was no way to form any opinion about the prospects for their favorability.

Tens of millions of tagged $B \rightarrow f_{\mathcal{CP}}$ decays would be required:

In 1980, the world's best source of B -mesons was CESR, with a production rate of $\sim 30 B\bar{B}$ events/day, of which only half were the desired $B^0\bar{B}^0$ pairs. Sanda's golden mode was

$B^0(\bar{B}^0) \rightarrow K_S J/\psi$, which he estimated to have an $\mathcal{O}(10^{-3})$ branching fraction, and this implied that the fractional probability of usable events would be

$$\mathcal{F}_{K_S J/\psi} < \underbrace{\mathcal{B}(B^0 \rightarrow K_S J/\psi)}_{\sim 10^{-3}} \underbrace{\mathcal{B}(K_S \rightarrow \pi^+ \pi^-)}_{\sim 10^{-1}} \underbrace{\mathcal{B}(J/\psi \rightarrow \ell\ell)}_{\sim 10^{-1}} (\epsilon_{\text{trk}})^4 \epsilon_{\text{eff}}^{\text{tag}} \approx 10^{-5}, \quad (40)$$

where ϵ_{trk} is the efficiency for charged track detection that, even in a nearly perfect detector, cannot be much higher than $\epsilon_{\text{trk}} \approx 0.85$, and $\epsilon_{\text{eff}}^{\text{tag}} \approx 0.3$ is an estimate of the maximum possible effective B -flavor tagging efficiency. Thus, an $\mathcal{A}_{KJ/\psi}^{\mathcal{CP}}$ measurement with even modest precision would require ~ 30 M $B\bar{B}$ events (and a million days of operation at 1980 state of the art collider and detector performance levels).

B^0 - \bar{B}^0 mixing had to be substantial, $|V_{cb}|$ had to be small, and $|V_{ub}|$ even smaller:

An essential part of the Sanda-Carter scheme is that the fraction of B^0 mesons that oscillate into a \bar{B}^0 before they decay has to be reasonably large. This meant that ΔM_B would have to be similar in magnitude to $\Gamma_B = 1/\tau_B$, which was then known to be $\Gamma_B \approx 4.4 \times 10^{-10}$ MeV. In the 1980s, when the t -quark mass was (almost universally) expected to be $m_t \sim 35$ GeV, calculations [54] found $\Delta M_B \approx 1.2 \times 10^{-10}$ MeV. If this were the case, only $\sim 6\%$ of the tagged B^0 -mesons would oscillate into a \bar{B}^0 before decaying. (After three lifetimes, the $\sin \Delta M_B \tau$ factor in eqn. 39 would have barely reached 0.5.) In addition the B lifetime had to be relatively long: *i.e.*, $|V_{cb}| < 0.1$, and $|V_{ub}| < |V_{cb}|$.

The time sequence between the tag- and $f_{\mathcal{CP}}$ -decay has to be distinguished:

The asymmetry in eqn. 39 has opposite signs for negative and positive values of τ , which makes it essential to distinguish between events in which the $B_{\mathcal{CP}}$ decays occurred first from those when it decayed last. The B mesons that are produced in $\Upsilon(4S) \rightarrow B\bar{B}$ decays have c.m. momenta $p_B = 327$ MeV/ c , corresponding to $\gamma\beta = 0.062$ and have a mean decay distance of $\beta\gamma c\tau_B = 28 \mu\text{m}$, which is unmeasurably small in a c.m. e^+e^- collider environment. For a collider operating at the $\Upsilon(4S)$, it would be impossible to distinguish the time sequence of the $B_{\mathcal{CP}}$ and B_{tag} decays.

Since the existence of six-quarks was pretty well established, the KM-mechanism provided a compelling and almost obvious mechanism for explaining the existence of \mathcal{CP} violation. However the prospects or a conclusive experimental test of this idea seemed hopeless. The fortuitous set of circumstances that made studies of \mathcal{CP} violation in the neutral kaon system possible seemed unlikely to be repeated.

5.3 Three miracles

Nevertheless, in spite of these obstacles, Sanda maintained a nearly mystical belief that “*Mother Nature has gone out of Her way to show us \mathcal{CP} violation, and She will also show us the way to the*

fundamental theory” [50], and forcefully advocated an aggressive program of experimental investigations of \mathcal{CP} violation in the decays of B mesons. However, for the reasons itemized above, Sanda’s advocacy was initially met with considerable skepticism from his colleagues in both the theoretical and experimental physics communities.

And then three miracles occurred:

Miracle 1: B^0 - \bar{B}^0 mixing was discovered at DESY:

The most exciting event in flavor physics during the 1980s was the 1987 discovery of a large signal for B^0 - \bar{B}^0 mixing by the ARGUS experiment at DESY [55]. The strength of the mixing was clear evidence that the top-quark mass, now known to be 173 GeV, was nearly an order of magnitude larger than expected, which was shocking news to almost everyone. This discovery, coupled with the 1.5 ps B -meson life-time measurements from PEP and PETRA that translated into $|V_{cb}| \approx 0.04$, and the suppression of $b \rightarrow u$ relative to $b \rightarrow c$ transitions meant that $|V_{ub}|$ was about a factor of ten smaller than $|V_{cb}|$. These measurements confirmed Sanda’s strong belief that Mother Nature would indeed help us “find the way to the fundamental theory.”

Miracle 2: Three-order-of-magnitude improvement in e^+e^- collider luminosity:

Advances in the understanding and modeling of beam dynamics, the use of separate magnet rings that enabled multibunch collisions, and major advances in RF feedback systems provided the huge increases in the $e^+e^- \rightarrow \Upsilon(4S)$ production rate that were required by the experiment [56, 57].

Miracle 3: The invention of asymmetric e^+e^- colliders and innovations in detector technology:

Pierre Oddone realized that an e^+e^- collider operating at the $\Upsilon(4S)$ resonance with a modest (*i.e.*, factor of ~ 2) difference between the e^+ and e^- beam energies would produce boosted B -mesons with $\mathcal{O}(100 \mu\text{m})$ separation distances between the B_{tag} and $B_{\mathcal{CP}}$ vertices [58]. This idea, coupled with the concurrent development of high resolution silicon-strip vertex detectors that could measure such small displacements in a collider environment [59–62], offered a realistic solution to the decay-time-sequence determination problem. Parallel improvements in detector performance levels in areas of particle identification [63, 64], and γ -ray & K_L detection [65–67] advanced the state of the art levels of detection efficiencies and \mathcal{B} -flavor-tagging quality.

6 First measurements of the KM angle β

At leading order, measurements of the eqn. 39 Carter-Sanda asymmetry determines $\sin 2\beta$, where $\beta = \tan^{-1}(\bar{\eta}/(1 - \bar{\rho}))$ and $\bar{\eta}$ & $\bar{\rho}$ are the modified Wolfenstein parameters described in Section 4.2.2. Prior to the summer of 2001, the best measurements of $\sin 2\beta$ were from CDF [68]

(0.79 ± 0.44), BaBar [69] (0.34 ± 0.21) and Belle [70] (0.58 ± 0.34). The BaBar result was based on a sample of 23 M $\Upsilon(4S) \rightarrow B\bar{B}$ events and Belle, which was struggling with electron cloud effects in the KEKB positron ring [71], had a smaller data sample of 11 M $\Upsilon(4S) \rightarrow B\bar{B}$ events. Each of the three measurements were about 1.5σ from zero, and their weighted average, 0.46 ± 0.17 indicated a non-zero \mathcal{CP} violation at the $\sim 2.5\sigma$ level. The situation was tantalizing, but not conclusive.

This changed in August 2001 when, in back-to-back articles in Physical Review Letters, BaBar, now with a sample of 32 M $B\bar{B}$ pairs, reported [72]

$$\sin 2\beta = 0.59 \pm 0.14 \pm 0.05 \quad \text{BaBar (2001)}, \quad (41)$$

and Belle, with a 31 M $B\bar{B}$ pair data sample, reported [73]

$$\sin 2\beta = 0.99 \pm 0.14 \pm 0.06 \quad \text{Belle (2001)}. \quad (42)$$

The BaBar data sample contained 803 $B_{\mathcal{CP}}$ event candidates with a signal purity of about 80%. The top three panels in Fig. 8a show the BaBar experiment's measured time distributions for $\xi_{\mathcal{CP}}=+1$ $B_{\mathcal{CP}}$ decays. The uppermost plot shows the number of B^0 tags where it is evident that there are more events with $\tau > 0$ than with $\tau < 0$, while the \bar{B}^0 tags in the panel beneath it display an opposite pattern. The third panel shows the bin-by-bin asymmetry where there is a clear indication of the sine-like behavior that is expected for a \mathcal{CP} violation as given in eqn. 39. The three lowest panels show the corresponding results for $B \rightarrow K_L J/\psi$ decays with $\xi_{\mathcal{CP}}=-1$, where the τ -dependent asymmetries have opposite signs, again as expected.

Figure 8b shows the results from the Belle experiment [73] that used 747 $\xi_{\mathcal{CP}}=-1$ B -decay candidates (mostly $B \rightarrow K_S J/\psi$) with 92% purity and 569 $\xi_{\mathcal{CP}}=+1$ $B \rightarrow K_L J/\psi$ decay candidates with 61% purity. The top plot on the left side of the panel shows the proper-time distribution for the $\xi_{\mathcal{CP}}=-1$ modes *minus* that for $\xi_{\mathcal{CP}}=+1$ modes. The 2nd and 3rd panels from the top show the $\xi_{\mathcal{CP}}=-1$ and $\xi_{\mathcal{CP}}=+1$ modes separately, where the different \mathcal{CP} modes have opposite-sign asymmetries as expected. The curves show the results of fits to the data. The bottom panel shows the asymmetry for a large sample of self-tagged (non- \mathcal{CP} eigenstate) B decays ($B^0 \rightarrow D^{(*)-} \pi^+$, $D^{*-} \rho^+$, $K^{*0}(K^+ \pi^-) J/\psi$, and $D^{*-} \ell^+ \nu$), where a non-zero asymmetry would have to be due to instrumental effects; the fit to this sample returned an asymmetry amplitude of 0.05 ± 0.04 .

The open circles in the plot on the right side of Fig. 8b show the time distribution for \bar{B}^0 -tags ($q=1$) in $\xi_{\mathcal{CP}}=-1$ $B_{\mathcal{CP}}$ decays plus B^0 -tags ($q=-1$) in $\xi_{\mathcal{CP}}=+1$ decays (*i.e.*, $q\xi_{\mathcal{CP}}=-1$), with the fit results shown as a dashed curve. The black dots and solid curve show the sum of the opposite combinations ($q\xi_{\mathcal{CP}}=+1$) and their fit result.

The BaBar and Belle results excluded a zero value for $\sin 2\beta$ by 4σ , and 6σ , respectively, and their combined significance established conclusively that \mathcal{CP} symmetry is violated in the $B-\bar{B}$

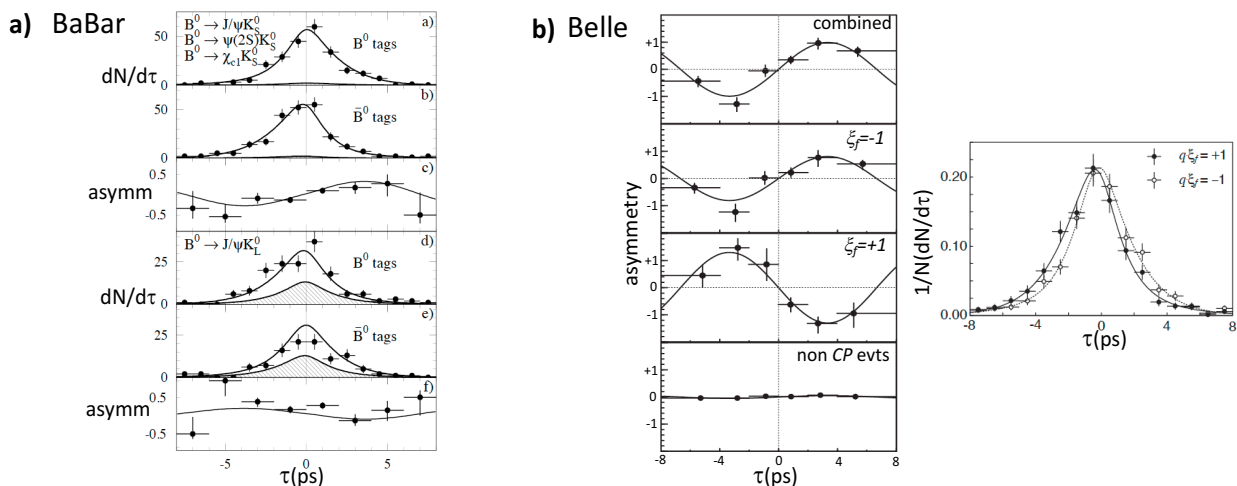


Fig. 8: **a)** BaBar results: the top three plots show the proper time distributions and fit results for B^0 -tags, \bar{B}^0 -tags and their asymmetry for $\xi_{CP}=-1$ B_{CP} decays. The bottom three plots show the same distributions for $\xi_{CP}=+1$ B_{CP} decays. The shaded areas indicate the background levels. (From ref. [72]). **b)** Belle results: the top three plots on the left show the time-dependent asymmetry and fit results for: (from top down) the combined ($\xi_{CP}=-1$ events *minus* $\xi_{CP}=+1$ events) samples; the $\xi_{CP}=-1$ (mostly $K_S J/\psi$) events; and the $\xi_{CP}=+1$ ($K_L J/\psi$) events. The bottom plot shows results for non- CP -eigenstate decay modes where no asymmetry is expected. The open circles and dashed curve on the right show background-subtracted time distributions for $q\xi_{CP}=-1$ events and the fit results. The solid circles and curve show the same quantities for $q\xi_{CP}=+1$ events. (From ref. [73]).

mixing process as as predicted by the KM six-quark model. Makoto Kobayashi and Toshihide Maskawa shared half of the 2008 Nobel Prize in Physics “for the discovery of the origin of the broken symmetry which predicts the existence of at least three families of quarks in nature.” The Nobel committee’s remarks that accompanied their announcement included the following:

“[Kobayashi and Maskawa] explained broken symmetry within the framework of the Standard Model, but required that Model be extended to three families of quarks. These predicted, hypothetical new quarks have recently appeared in physics experiments. As late as 2001, the two particle detectors BaBar at Stanford, USA and Belle at Tsukuba, Japan, both detected broken symmetries independently of each other. The results were exactly as Kobayashi and Maskawa had predicted almost three decades earlier.”

During the two decades following the Belle and BaBar reports, there have been many hundreds of measurements of non-zero CP violating symmetries in B meson decays, mostly by BaBar and Belle, which continued operating until 2008 and 2010, respectively. This program is being continued by the LHCb [74], an experiment specialized for heavy flavor physics at the CERN Large Hadron Collider that began operating in 2010, and Belle II [75], an upgraded version of Belle at KEK that began operating in 2018. As is discussed in more detail in other contributions to this symposium, all of these hundreds of measured CP violations are well explained as being due to the effects of the single KM CP -phase angle δ .

The KM model has been a remarkable success.

7 A few comments on mixing and \mathcal{CP} violation in the neutrino sector

As mentioned above, the two-doublet nature of the of the leptons was identified in 1961 [16], a decade before it was established for quarks. The notion of neutrino-mixing was first proposed four years earlier by Pontecorvo [76] in 1957 and the PMNS neutrino mixing matrix was suggested Maki, Nakagawa and Sakata [77] in 1962, a year before Cabibbo’s paper appeared. When the τ -lepton was discovered in 1975, the six-lepton picture was established and the PMNS matrix for neutrinos expanded to the same 3×3 structure as the CKM matrix for quarks. If neutrinos are Dirac particles, the mathematics of the neutrino-flavor mixing matrix are the same as the KM matrix with three mixing angles θ_{ij} and one \mathcal{CP} violating phase $\delta_{\mathcal{CP}}$, the so-called the “Dirac” phase. If neutrinos are Majorana particles, lepton number is not conserved and the matrix’s number of degrees of freedom increases by two and there are two additional \mathcal{CP} -violating “Majorana” phases [78] that have no measurable effects on neutrino mixing experiments (which makes them hard to measure). The commonly used parameterization of the PMNS matrix that doesn’t include the Majorana phases uses same mixing angle definitions as the eqn. 26) Chau-Keung version of the CKM matrix, but with the sequence of rotations reversed, *i.e.*,

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & e^{-i\delta_{\mathcal{CP}}} s_{13} \\ 0 & 1 & 0 \\ -e^{-i\delta_{\mathcal{CP}}} s_{13} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \quad (43)$$

where ν_1, ν_2, ν_3 denote the three neutrino mass eigenstates and θ_{12}, θ_{23} , and θ_{13} are the “solar,” “atmospheric,” and “reactor” neutrino mixing angles. The explicit form of the matrix is:¹⁸

$$\begin{aligned} U_{\text{PMNS}}^{\text{Dirac}} &= \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{\mathcal{CP}}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{\mathcal{CP}}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{\mathcal{CP}}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{\mathcal{CP}}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{\mathcal{CP}}} & c_{23}c_{13} \end{pmatrix}, \end{aligned} \quad (44)$$

¹⁸ \mathcal{CP} -violating Majorana phases ψ_1 & ψ_2 can be included by $U_{\text{PMNS}}^{\text{Maj}} = U_{\text{PMNS}}^{\text{Dirac}} \mathbb{P}$, where $\mathbb{P} = \begin{pmatrix} e^{i\psi_1} & 0 & 0 \\ 0 & e^{i\psi_2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

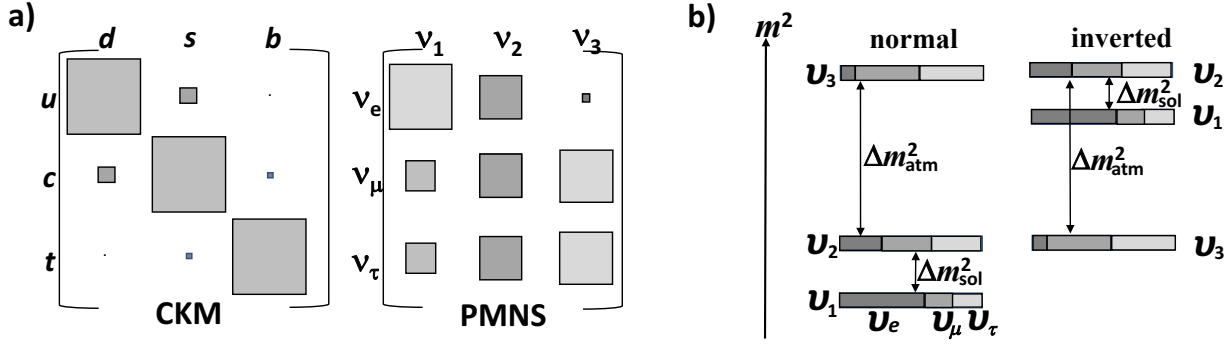


Fig. 9: **a)** The area of the squares illustrates the magnitudes of the CKM (*left*) and PMNS (*right*) matrix elements. **b)** A not-to-scale illustration of the normal and (*left*) and inverted (*right*) neutrino mass hierarchy. The flavor content of each of the ν_1, ν_2, ν_3 mass eigenstate is indicated by different shades of gray.

and the PDG 2020 [38] world averages for the three rotation angles are:

$$\begin{aligned}
 \sin^2 \theta_{12} &= 0.307 \pm 0.013 \quad \Rightarrow \theta_{12} = 33.6^\circ \pm 0.8^\circ & (45) \\
 \sin^2 \theta_{23} &= 0.545 \pm 0.021 \quad \Rightarrow \theta_{23} = 47.6^\circ \pm 1.2^\circ \\
 \sin^2 \theta_{13} &= 0.0218 \pm 0.007 \quad \Rightarrow \theta_{13} = 8.48^\circ \pm 0.14^\circ.
 \end{aligned}$$

Although mixing in the quark- and lepton-sectors have the same mathematical structure, there are important differences in their practical applications, including:

Strikingly different hierarchies: The three PMNS mixing angles listed above differ in values by at most a factor of five, in contrast with the the CKM-matrix where the corresponding mixing angles are smaller and span a two-order-of-magnitude range in magnitudes:

$$\theta_{12}^{\text{CKM}} \approx 13.0^\circ \quad \theta_{23}^{\text{CKM}} \approx 2.4^\circ \quad \theta_{13}^{\text{CKM}} \approx 0.20^\circ. \quad (46)$$

Differences between the two matrices are illustrated in Fig. 9a, where the areas of the squares are proportional to the magnitudes of the matrix elements. Here Nature has been helpful again; if the PMNS matrix had a hierarchy that was similar to that of the CKM matrix, neutrino oscillations and neutrino masses would likely not have been discovered.

They operate in opposite “directions:” In the CKM case, the quark mass-eigenstates are well known and the matrix is used to determine the flavor states. For the PMNS case, the neutrino flavor states are well known and the matrix is used to determine the mass eigenstates. Oscillation experiments have determined the mass-difference hierarchies shown in Fig. 9b, where [38]

$$\text{atmos} : \Delta m_{32}^2 = \pm(2.44 \pm 0.03) \times 10^{-3} (\text{eV})^2 \quad \text{and} \quad \text{solar} : \Delta m_{21}^2 = (7.53 \pm 0.18) \times 10^{-5} (\text{eV})^2.$$

The still unknown sign of $\Delta m_{32}^2 (= m_3^2 - m_2^2)$ leaves two possible hierarchies as shown in Fig. 9b.

Quarks decay, neutrinos (probably) don't: CKM-related measurements are almost entirely based on measurements of decay processes with a formalism for oscillations and \mathcal{CP} violations is done in the hadron's restframe. In contrast case, PMNS-related measurement always involve production processes that produce pure flavor states and detection experiments located at some baseline distance that determine how the neutrino flavor-contents changed during their propagation to the detector. Since the neutrino's restframe is ill defined,¹⁹ the formalism is usually done in the laboratory frame and expressed in terms of E_ν - and Δm^2 -dependent oscillation lengths, *i.e.*,

$$l(E_\nu) \equiv \lambda/2\pi = E_\nu/1.27\Delta m^2, \quad (47)$$

where the factor of 1.27 is specific to $(l, E_\nu, \Delta m^2)$ units of (m, MeV, eV²), or (km, GeV, eV²). For the atmospheric and solar mass differences given above, these are

$$l^{\text{atm}}(E_\nu) \approx 320 \text{ (km/GeV)} \times E_\nu \quad \text{and} \quad l^{\text{sol}}(E_\nu) \approx 10,500 \text{ (m/MeV)} \times E_\nu,$$

where the units km/GeV and m/MeV are equivalent (and interchangeable), with km/GeV usually attached to atmospheric and m/MeV to solar for historical reasons.

The 320 km atmospheric length is long, but not impossibly long. It is nearly the same as the 295 km distance between J-PARC and Super Kamiokande (and the soon-to-be-completed HyperK detector), and corresponds to a 90° phase-change-induced oscillation maximum for 600 MeV muon-neutrinos that can be copiously produced by the J-PARC synchrotron. The 10.5 km solar length corresponds to an oscillation maximum for $E_\nu = 3$ MeV reactor anti-electron-neutrinos at a baseline of 50 km, which is the baseline for the JUNO reactor neutrino experiment that will soon be operating in China. The HyperK [79] and JUNO [80] experiments will certainly not be easy, but they will be done. If the neutrino oscillation lengths were factors of two or more longer, both experiments would be much more difficult, if not impossible.

8 Conclusions

The related subjects of \mathcal{CP} -violation and quark & neutrino flavor mixing provide peeks into some of Nature's most intimate secrets. When Fitch and collaborators measured the $\sim 0.2\%$ branching fraction for $K_L \rightarrow \pi^+ \pi^-$ in 1963, they were seeing the influence of the t -quark that wasn't discovered until thirty years later (see Fig. 4). Thanks to Kobayashi and Maskawa, the existence and most of the properties of the t -quark (other than its mass) were pretty well understood more than twenty years before it was discovered.

¹⁹ If the lightest neutrino has zero mass it doesn't have a restframe.

8.1 Nature has been kind

(Einstein famously said “The Lord God is subtle, but not malicious.”)

A common thread that characterizes this entire story is that we have been able to probe these subjects at considerable depth, even though it didn't *a priori* have to be that way. The 0.2% $K_L \rightarrow \pi^+ \pi^-$ branching fraction is as large as it is because of the phase-space suppression of the partial decay width for the \mathcal{CP} -allowed $K_L \rightarrow 3\pi$ mode that has a Q -value of only 83 MeV. This is a consequence of the relative masses of the K - and π -mesons: if the kaon mass were higher and/or the pion mass were lighter, the $K_L \rightarrow 3\pi$ partial width would be larger, and the $K_L \rightarrow \pi^+ \pi^-$ branching fraction would be reduced. It wouldn't take very large changes in these masses to push the $\pi^+ \pi^-$ branching fraction down to a value that was below the Fitch experiment's sensitivity level. Unlike parity violation, which is a huge effect that is difficult to miss, the particle physics consequences of \mathcal{CP} -violations in processes other than K_L decays have only been seen in elaborate, highly focused experiments that, absent the K_L measurements, very likely wouldn't have occurred.

As described above, the KM phase was only measurable because Nature's parameters are aligned in a way that meet the stringent requirements that were first identified in the Carter-Sanda papers, including a strong $|V_{us}| > |V_{cb}| > |V_{ub}|$ hierarchy and a large enough t -quark mass to produce a B - \bar{B} mixing rate that is close to the B -meson lifetime, but not so large that the mixing rate was much faster than the lifetime. As a result, the phase could be measured, as could all three of the CKM matrix's mixing angles.

Remarkably, the same thread extends into the neutrino sector in which the PMNS matrix has a very different, almost flat hierarchy that facilitated the discovery of neutrino oscillations. This hierarchy has also enabled precise measurements of all three of its mixing angles and will likely allow for a determination of its Dirac phase in the not-so-distant future. Moreover, and as mentioned above, Nature's choices of the differences between the ν_1, ν_2, ν_3 eigenstate masses have made these measurements realizable.

8.2 What's next?... \mathcal{CPT} ?

Although \mathcal{CP} violations only show up as small, subtle effects in particle physics experiments, the influence of \mathcal{CP} violations on the evolution of the Universe is glaringly obvious [81]. However, the \mathcal{CP} violations that we measure in K -, B - and (recently [82]) D -meson decays that are produced by the KM mechanism for quarks and, likely, leptons, cannot nearly account for the matter-antimatter asymmetry of the Universe (see, *e.g.*, ref. [83]). There must be other sources of \mathcal{CP} violation that have yet to be discovered, and these are the primary motivations for the LHCb and Belle II experiments. With the KM phase, Nature has given us a glimpse of \mathcal{CP} violation, but not the whole story.

But what about CPT ? The CPT theorem [84] states that any quantum field theory that is *Lorentz invariant*, has *local point-like interaction vertices*, and is *Hermitian* (i.e., conserves probability) is invariant under the combined operations of \mathcal{C} , \mathcal{P} and \mathcal{T} . Since the three QFTs that make up the Standard Model—QED, QCD, and Electroweak theory—all satisfy these criteria, CPT symmetry has been elevated to a kind of hallowed status in particle physics. But the non-renormalizability of quantum gravity calls into question the validity of the locality assumption [85], and suggests that at some scale, CPT has to be violated. Strictly speaking, this violation only has to occur at impossibly high energies near the $\mathcal{O}(10^{19} \text{ GeV})$ Planck scale, but maybe the same thread of Nature that gives us a taste of \mathcal{CP} violations at energy scales well below the scale needed to explain the baryon asymmetry of the Universe, will give us a hint of CPT violation at scale below the one that’s needed to rescue quantum gravity. In any case, since it is a fundamental feature of the Standard Model, CPT invariance should be routinely challenged at the highest feasible experimental sensitivities.

To date the most stringent test of the CPT prediction that particle-antiparticle masses are equal comes from kaon physics²⁰ [86, 87] and sets a 90% C.L. limit on the K^0 - \bar{K}^0 mass difference of

$$|M_{\bar{K}^0} - M_{K^0}| < 5 \times 10^{-19} \text{ GeV}, \quad (48)$$

which is seven orders-of-magnitude more stringent than that for $m_{\bar{e}} - m_e$ and nine orders-of-magnitude more stringent than the $m_{\bar{p}} - m_p$ limit. This high sensitivity is because of the magic of the virtual processes in the Fig. 4a box diagram and the unique properties of the neutral kaons.

The kaon result is based on experiments done nearly thirty years ago with data samples containing tens of millions of $K \rightarrow \pi^+ \pi^-$ decays. One of the reasons these measurements have not been updated since then is that technologies for improved sources of flavor-tagged neutral kaons have not been pursued. However, if the above-described three order of magnitude improvements in $e^+ e^-$ collider luminosity that were developed for the B -factories would be applied to a dedicated collider operating at J/ψ mass peak, multi-billion-event/year samples of flavor-tagged neutral kaons, produced via $J/\psi \rightarrow K^\mp \pi^\pm K^0 (\bar{K}^0)$ decays, would be available to support CPT tests with more than an order-of-magnitude improved sensitivity. More modest improvements in CPT sensitivity would be provided by new colliders in the τ -charm energy range that are being proposed in China [88] and Russia [89], if they spend sufficient time operating at the J/ψ peak.

Maybe during the next sixty years, CPT violation studies will prove to be as interesting and provocative as \mathcal{CP} studies have been during the past sixty years.

²⁰ The CPT test in kaon physics involves a comparison the phase of the η_{+-} \mathcal{CP} violation parameter in $K_L \rightarrow \pi^+ \pi^-$ decays with the “superweak phase” $\phi_{\text{SW}} \equiv \tan^{-1} (2(M_{K_L} - M_{K_S}) / (\Gamma_{K_S} - \Gamma_{K_L}))$.

Acknowledgment

I thank the organizers of KM50 for inviting me to participate in this interesting symposium, the editors of PTEP for inviting this manuscript, and this paper's referee, whomever he or she may be, for many important corrections and helpful suggestions. This work was supported in part by the National Research Foundation of Korea under Contract No. NRF-2022R1A2C1092335.

References

- [1] R. P. Feynman and Murray Gell-Mann, *Phys. Rev.*, **109**, 193–198 (1958).
- [2] Nicola Cabibbo, *Phys. Rev. Lett.*, **10**, 531–533 (1963).
- [3] Byron P. Roe, Daniel Sinclair, John L. Brown, Donald A. Glaser, John A. Kadyk, and George H. Trilling, *Phys. Rev. Lett.*, **7**, 346–348 (1961).
- [4] Murray Gell-Mann and M Levy, *Nuovo Cim.*, **16**, 705 (1960).
- [5] F. Eisele, R. Engelmann, H. Filthuth, W. Foehlisch, V. Hepp, E. Leitner, W. Presser, H. Schneider, and Gunter Zech, *Z. Phys.*, **225**, 383–392 (1969).
- [6] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Phys. Rev. D*, **2**, 1285–1292 (1970).
- [7] J. B. Gerhart, *Phys. Rev.*, **95**, 288–289 (1954).
- [8] D. L. Hendrie and J. B. Gerhart, *Phys. Rev.*, **121**, 846–849 (1961).
- [9] Murray Gell-Mann, *Phys. Lett.*, **8**, 214–215 (1964).
- [10] G. Zweig, An SU(3) model for strong interaction symmetry and its breaking. Version 1, CERN-TH-401, unpublished (1964).
- [11] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Phys. Rev. Lett.*, **13**, 138–140 (1964).
- [12] L. B. Leipuner, W. Chinowsky, R. Crittenden, R. Adair, B. Musgrave, and F. T. Shively, *Phys. Rev.*, **132**, 2285–2290 (1963).
- [13] Makoto Kobayashi and Toshihide Maskawa, *Prog. Theor. Phys.*, **49**, 652–657 (1973).
- [14] Kiyoshi Niu, Eiko Mikumo, and Yasuko Maeda, *Prog. Theor. Phys.*, **46**, 1644–1646 (1971).
- [15] Toshihide Maskawa, *Rev. Mod. Phys.*, **81**, 1027 (2009).
- [16] G. Danby, J. M. Gaillard, Konstantin A. Goulianos, L. M. Lederman, Nari B. Mistry, M. Schwartz, and J. Steinberger, *Phys. Rev. Lett.*, **9**, 36–44 (1962).
- [17] J. D. Bjorken and S. L. Glashow, *Phys. Lett.*, **11**, 255–257 (1964).
- [18] Pekka Tarjanne and Vigdor L. Teplitz, *Phys. Rev. Lett.*, **11**, 447–448 (1963).
- [19] Ziro Maki, *Prog. Theor. Phys.*, **31**, 331 (1964).
- [20] Yasuo Hara, *Phys. Rev.*, **134**, B701–B704 (1964).
- [21] L. Maiani, *Phys. Lett. B*, **62**, 183–186 (1976).
- [22] Martin L. Perl et al., *Phys. Rev. Lett.*, **35**, 1489–1492 (1975).
- [23] S. W. Herb et al., *Phys. Rev. Lett.*, **39**, 252–255 (1977).
- [24] F. Abe et al., *Phys. Rev. Lett.*, **74**, 2626–2631 (1995), hep-ex/9503002.
- [25] S. Abachi et al., *Phys. Rev. Lett.*, **74**, 2632–2637 (1995), hep-ex/9503003.
- [26] Vernon D. Barger, Wai-Yee Keung, and R. J. N. Phillips, *Phys. Rev. D*, **24**, 1328, The journal editors revised the title in the published version (1981).
- [27] G. C. Branco and Hans Peter Nilles, *Nucl. Phys. B*, **151**, 529–535 (1979).
- [28] M. Gorn and E. A. Paschos, *Phys. Lett. B*, **101**, 255–258 (1981).
- [29] Utpal Sarkar and Amitava Raychaudhuri, DOES BARE BOTTOM RULE OUT THE TOPLESS E6 MODEL? (4 1983).
- [30] Y. Yamanaka, H. Sugawara, and S. Pakvasa, *Phys. Rev. D*, **25**, 1895, [Erratum: *Phys. Rev. D* 29, 2135 (1984)] (1982).
- [31] Sandip Pakvasa and Hirotaka Sugawara, *Phys. Rev. D*, **14**, 305 (1976).
- [32] Benjamin W. Lee, *Phys. Rev. D*, **15**, 3394 (1977).
- [33] John R. Ellis, M. K. Gaillard, Dimitri V. Nanopoulos, and S. Rudaz, *Nucl. Phys. B*, **131**, 285, [Erratum: *Nucl. Phys. B* 132, 541 (1978)] (1977).
- [34] Steven Weinberg, *Experientia Suppl.*, **31**, 339–352 (1977).
- [35] Guido Altarelli, L. Baulieu, N. Cabibbo, L. Maiani, and R. Petronzio, *Nucl. Phys. B*, **125**, 285, [Erratum: *Nucl. Phys. B* 130, 516 (1977)] (1977).
- [36] Haim Harari and Miriam Leurer, *Phys. Lett. B*, **181**, 123–128 (1986).
- [37] E. Fernandez et al., *Phys. Rev. Lett.*, **51**, 1022 (1983).
- [38] P. A. Zyla et al., *PTEP*, **2020**(8), 083C01 (2020).
- [39] N. Lockyer et al., *Phys. Rev. Lett.*, **51**, 1316 (1983).
- [40] D. E. Klem et al., *Phys. Rev. Lett.*, **53**, 1873 (1984).

- [41] M. Althoff et al., Phys. Lett. B, **149**, 524–532 (1984).
- [42] A. Chen et al., Phys. Rev. Lett., **52**, 1084 (1984).
- [43] Ling-Lie Chau and Wai-Yee Keung, Phys. Rev. Lett., **53**, 1802 (1984).
- [44] Lincoln Wolfenstein, Phys. Rev. Lett., **51**, 1945 (1983).
- [45] M. Gronau and J. Schechter, Phys. Rev. D, **31**, 1668–1675 (1985).
- [46] M. Aguilar-Benitez et al., Phys. Lett. B, **170**, 1–350 (1986).
- [47] Andrzej J. Buras, Markus E. Lautenbacher, and Gaby Ostermaier, Phys. Rev. D, **50**, 3433–3446 (1994), hep-ph/9403384.
- [48] J. Charles, Andreas Hocker, H. Lacker, S. Laplace, F. R. Le Diberder, J. Malcles, J. Ocariz, M. Pivk, and L. Roos, Eur. Phys. J. C, **41**(1), 1–131, [updated results and plots available at: <http://ckmfitter.in2p3.fr>] (2005), hep-ph/0406184.
- [49] M. Kobayashi, Prog. Theor. Phys., **92**, 287–288 (1994).
- [50] A. I. Sanda, Prog. Theor. Phys., **122**, 103–123 (2009).
- [51] D. Andrews et al., Phys. Rev. Lett., **45**, 219 (1980).
- [52] Ashton B. Carter and A. I. Sanda, Phys. Rev. Lett., **45**, 952 (1980).
- [53] Ashton B. Carter and A. I. Sanda, Phys. Rev. D, **23**, 1567 (1981).
- [54] A. J. Buras, W. Slominski, and H. Steger, Nucl. Phys. B, **245**, 369–398 (1984).
- [55] H. Albrecht et al., Phys. Lett. B, **192**, 245–252 (1987).
- [56] Y. Funakoshi, Operational experience with crab cavities at KEKB, In *ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders*, pages 27–36 (2014), arXiv:1410.4036.
- [57] J. Seeman et al., ICFA Beam Dyn. Newslett., **48**, 72–87 (2009).
- [58] P. Oddone, DETECTOR CONSIDERATIONS, In Donald H. Stork, editor, *Proceedings, Conference on Linear Collider B-B Factory Conceptual Design, Los Angeles, California, 26-30 Jan. 1987*, pages 423–446, Los Angeles, Calif (1987). World Scientific.
- [59] A. Litke et al., Nucl. Instrum. Meth. A, **265**, 93–98 (1988).
- [60] G. Batignani et al., Nucl. Phys. B Proc. Suppl., **23**, 291–296 (1991).
- [61] C. Bozzi et al., Nucl. Instrum. Meth. A, **447**, 15–25 (2000).
- [62] G. Alimonti et al., Nucl. Instrum. Meth. A, **453**, 71–77 (2000).
- [63] I. Adam et al., Nucl. Instrum. Meth. A, **433**, 121–127 (1999).
- [64] T. Iijima et al., Nucl. Instrum. Meth. A, **453**, 321–325 (2000).
- [65] Bernard Aubert et al., Nucl. Instrum. Meth. A, **479**, 1–116 (2002), hep-ex/0105044.
- [66] A. Abashian et al., Nucl. Instrum. Meth. A, **479**, 117–232 (2002).
- [67] A. Abashian et al., Nucl. Instrum. Meth. A, **449**, 112–124 (2000).
- [68] T. Affolder et al., Phys. Rev. D, **61**, 072005 (2000), hep-ex/9909003.
- [69] Bernard Aubert et al., Phys. Rev. Lett., **86**, 2515–2522 (2001), hep-ex/0102030.
- [70] A. Abashian et al., Phys. Rev. Lett., **86**, 2509–2514 (2001), hep-ex/0102018.
- [71] L. F. Wang, H. Fukuma, K. Ohmi, S. Kurokawa, K. Oide, and F. Zimmermann, Phys. Rev. ST Accel. Beams, **5**, 124402 (2002).
- [72] Bernard Aubert et al., Phys. Rev. Lett., **87**, 091801 (2001), hep-ex/0107013.
- [73] Kazuo Abe et al., Phys. Rev. Lett., **87**, 091802 (2001), hep-ex/0107061.
- [74] I. Belyaev, G. Carboni, N. Harnew, C. Matteuzzi, and F. Teubert, Eur. Phys. J. H, **46**(1), 3 (2021), arXiv:2101.05331.
- [75] T. Abe et al., Belle II Technical Design Report (11 2010), arXiv:1011.0352.
- [76] B. Pontecorvo, Zh. Eksp. Teor. Fiz., **34**, 247 (1957).
- [77] Ziro Maki, Masami Nakagawa, and Shoichi Sakata, Prog. Theor. Phys., **28**, 870–880 (1962).
- [78] J. Schechter and J. W. F. Valle, Phys. Rev. D, **22**, 2227 (1980).
- [79] K. Abe et al., Hyper-Kamiokande Design Report (5 2018), arXiv:1805.04163.
- [80] Zelimir Djurcic et al., JUNO Conceptual Design Report (8 2015), arXiv:1508.07166.
- [81] A. D. Sakharov, J. Exptl. Theoret. Phys., **5**, 24–27 (1967).
- [82] Roel Aaij et al., Phys. Rev. Lett., **122**(21), 211803 (2019), arXiv:1903.08726.
- [83] Laurent Canetti, Marco Drewes, and Mikhail Shaposhnikov, New J. Phys., **14**, 095012 (2012), arXiv:1204.4186.
- [84] Julian S. Schwinger, Phys. Rev., **82**, 914–927 (1951).
- [85] Steven Weinberg and Edward Witten, Phys. Lett. B, **96**, 59–62 (1980).
- [86] B. Schwingerheuer et al., Phys. Rev. Lett., **74**, 4376–4379 (1995).
- [87] A. Apostolakis et al., Phys. Lett. B, **458**, 545–552 (1999).
- [88] M. Achasov et al., STCF Conceptual Design Report: Volume I - Physics & Detector (3 2023), arXiv:2303.15790.
- [89] A. Yu. Barnyakov, J. Phys. Conf. Ser., **1561**(1), 012004 (2020).