1. INTRODUCTION

As is well known, the flavor sector of the Standard Model (SM) violates CP, the combination of the discrete symmetries C and P. This suggests that flavor and CP violation have a common origin. The question of flavor concerns the fact that the SM fermions come in three families that are only distinguished by their masses. SU(2)l interactions lead to transitions between these families, which are governed by the mixing parameters in the CKM and PMNS matrices. These mixing parameters are completely unexplained in the SM. Furthermore, CP violation manifests in the SM through the non–zero phase $\delta_\text{CP}$ in the CKM matrix [1]. In the lepton sector, the latest measurements from T2K as well as the global fit for neutrino oscillation parameters also hint at non–zero value for the Dirac phase $\delta_L$ in the PMNS matrix [2], which will, if proved, establish violation of CP in the lepton sector.

The observed repetition of families, i.e., the fact that the quarks and leptons appear in 3 generations, hints at a flavor symmetry under which the generations transform nontrivially. The main punchline of this review is the statement that certain flavor symmetries clash with CP [3, 4]. In other words, CP violation can be entirely group theoretical in origin.

1.1. What is a physical CP transformation?

Charge conjugation C inverts, by definition, all currents. This implies that Standard Model representations R get mapped to their conjugates, $\overline{R}$. Likewise, parity P exchanges the $(0,1/2)$ and $(1/2,0)$ representations of the Lorentz group, which corresponds to complex conjugation at the level of $\text{SL}(2,C)$. That is, at the level of $G_{\text{SM}} \times \text{SL}(2,C)$, CP is represented by the (unique) nontrivial outer automorphism.

This fact has led to the suspicion that any nontrivial outer automorphism can be used to coin a valid CP transformation [5]. However, this turns out not to be the case [4]. To see this, let us review why we care about whether or not CP is violated. One reason we care is that CP violation is a prerequisite for baryogenesis [6], i.e., the creation of the matter–antimatter asymmetry of our universe. Therefore, a physical CP transformation exchanges particles and antiparticles, a requirement an arbitrary outer automorphism may or may not fulfill. As discussed in detail in [4], CP transformations are linked to class–inverting outer automorphisms.

1.2. CP and Clebsch–Gordan coefficients

It turns out that some finite groups do not have such outer automorphisms but still complex representations. These groups thus clash with CP! Further, they have no basis in which all Clebsch–Gordan coefficients (CGs) are real, and CP violation can thus be linked to the complexity of the CGs [3].

2. CP VIOLATION FROM FINITE GROUPS

2.1. The canonical CP transformation

Let us start by collecting some basic facts. Consider a scalar field operator

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3} \left[ a(p) e^{-i p \cdot x} + b^\dagger(p) e^{i p \cdot x} \right],$$

where $a$ annihilates a particle and $b^\dagger$ creates an antiparticle. The CP operation exchanges particles and antiparticles,

$$\begin{align*}
(CP)^{-1} a(p) CP &= \eta_{CP} b(-p), \\
(CP)^{-1} a^\dagger(p) CP &= \eta^*_{CP} b^\dagger(-p), \\
(CP)^{-1} b(p) CP &= \eta_{CP} a(-p), \\
(CP)^{-1} b^\dagger(p) CP &= \eta^*_{CP} a^\dagger(-p),
\end{align*}$$

where $\eta_{CP}$ is a phase factor. On the scalar fields, CP transformations act as

$$\phi(x) \xrightarrow{CP} \eta_{CP} \phi^* (P x).$$

At this level, $\eta_{CP}$ can be viewed as the freedom of rephasing the field, i.e., a choice of field basis. Later, when we replace $\eta_{CP}$ by some matrix $U_{CP}$, this will still reflect the freedom to choose a basis. The important message here is that there is a well–defined operation, the CP transformation, which exchanges particles with antiparticles. It is this very transformation which is broken in the $K^0 - \overline{K^0}$ system, and whose violation is a prerequisite for baryogenesis.

2.2. CP vs. outer automorphisms

Next let us review what CP does in the context of most of the continuous (i.e., Lie) groups. If the representation under consideration is real, the canonical CP does the job. For complex representations, CP involves a nontrivial outer automorphism (cf. Figure 1).

In particular, in the context of the Standard Model gauge group and the usual theories of grand unification (GUTs),

$$G_{\text{SM}} = \text{SU}(3)_C \times \text{SU}(2)_L \times U(1)_Y \subset \text{SU}(5) \subset \text{SO}(10) \subset E_6$$
The canonical $\mathcal{CP}$ transformation is not an (outer) automorphism of $T$ (in this basis). Therefore, in order to warrant $\mathcal{CP}$ conservation, one needs to impose a so-called generalized $\mathcal{CP}$ transformation $\tilde{\mathcal{CP}}$ under which $\phi \xrightarrow{\tilde{\mathcal{CP}}} \phi^*$ as usual but

$$
\begin{pmatrix}
  x_1 \\
  x_2 \\
  x_3 \\
  x_4
\end{pmatrix}
\xrightarrow{\tilde{\mathcal{CP}}}
\begin{pmatrix}
  x_1' \\
  x_2' \\
  x_3' \\
  x_4'
\end{pmatrix},
\begin{pmatrix}
  y_1 \\
  y_2 \\
  y_3 \\
  y_4
\end{pmatrix}
\xrightarrow{\tilde{\mathcal{CP}}}
\begin{pmatrix}
  y_1' \\
  y_2' \\
  y_3' \\
  y_4'
\end{pmatrix}.
\tag{2.7}
$$

2.4. Constraints on generalized $\mathcal{CP}$ transformations

In order for a $\mathcal{CP}$ transformation not to clash with the group, i.e., in order to avoid mapping something that is invariant under the symmetry transformations to something that isn’t (cf. (2.6)), it has to be an automorphism $u : G \to G$ of the group. An automorphism $u$ corresponding to a physical $\mathcal{CP}$ transformation has to fulfill the consistency condition [5] (see also [7])

$$
\rho(u(g)) = U_{\mathcal{CP}} \rho(g) U_{\mathcal{CP}}^\dagger \quad \forall g \in G.
\tag{2.8}
$$

Here, $U_{\mathcal{CP}}$ is a unitary matrix that enters the generalized $\mathcal{CP}$ transformation,

$$
\Phi(x) \xrightarrow{\mathcal{CP}} U_{\mathcal{CP}} \Phi'(P x),
\tag{2.9}
$$

where $\Phi$ denotes collectively the fields of the theory/model, and $P (t, \vec{x}) = (t, -\vec{x})$ as usual. In particular, each representation gets mapped on its own conjugate, i.e., $U_{\mathcal{CP}}$ is block-diagonal in Equation (2.9),

$$
\begin{pmatrix}
  \phi_{r_1} \\
  \phi_{r_2} \\
  \vdots
\end{pmatrix}
\xrightarrow{\mathcal{CP}}
\begin{pmatrix}
  U_{r_1} & & \\
  & U_{r_2} & \\
  & & \ddots
\end{pmatrix}
\begin{pmatrix}
  \phi_{r_1}' \\
  \phi_{r_2}' \\
  \vdots
\end{pmatrix},
\tag{2.10}
$$

where the $U_{r_i}$ are unitary matrices that depend on the representation $r_i$ only. The $a$ subscripts in $\phi_{r_a}$ label the particles whereas the $i_a$ subscripts indicate the representations, i.e., different particles can furnish the same representations under $G$. The transformation law (2.10) disagrees with [5], where it was suggested that one can use any outer automorphism in order to define a viable $\mathcal{CP}$ transformation.

Therefore, the requirement that the candidate transformation is a physical $\mathcal{CP}$ transformation, which exchanges particles and their antiparticles, amounts to demanding that $u$ be class-inverting. In all known cases, $u$ can be taken to be an automorphism of order two. Of course, this does not exclude the interesting possibility to make $\mathcal{CP}$ part of a higher–order transformation [8].

2.5. $\mathcal{CP}$ vs. $\mathcal{CP}$–like transformations

However, it is important to distinguish physical $\mathcal{CP}$ transformations, and their proper generalizations, from $\mathcal{CP}$–like transformations. Unfortunately, the latter have sometimes been called “generalized $\mathcal{CP}$ transformations” in the literature.\footnote{More detailed comments on the literature can be found in Section VII.15 of [9].}
However, some of the proposed “generalized CP transformations” do not warrant physical CP conservation. Thus, they do not have a connection to the observed CP violation in the CKM sector, nor to baryogenesis and so on. That is, the violation of physical CP is a prerequisite of a nontrivial decay asymmetry, but the violation of a so-called “generalized CP transformation” is not. That is to say some of the operations dubbed “generalized CP transformation” in the literature are not physical CP transformations, which is why we refer to them as “CP-like”.

Given all these considerations, it is a valid question whether or not one can impose a physical CP in any model. As mentioned above, this is not the case. Certain finite symmetries clash with CP. Here “clash” means that any physical CP transformation maps some G-invariant term(s) on non-invariant terms, and thus does not comply with G, i.e., is not an automorphism thereof. We will discuss next how one may tell those symmetries that clash with CP apart from those which do not.

2.6. The Bickerstaff–Damhus automorphism (BDA)

In a more group–theoretical language the question whether or not one can impose CP can be rephrased as whether or not a given finite group has a so-called Bickerstaff–Damhus automorphism (BDA) [10] u,\n
$$\rho_r(u(g)) = U_r \rho_r(g)^* U_r^+ \quad \forall g \in G \text{ and } \forall i,$$  \hspace{1cm} (2.11)\n
where U is unitary and symmetric. The existence of a BDA implies the existence of a basis in which all Clebsch–Gordan (CG) coefficients are real. In physics, such a basis is often referred to as “CP basis”. The connection between the BDA, the complexity of the CG’s, and CP has first been pointed out in [3].

Of course, this raises the question whether or not one can tell if a given group has an BDA. There is a rather simple criterion for this, based on the so-called extended twisted Frobenius–Schur indicator\(^2\) (see [10, 12] for the so-called twisted Frobenius–Schur indicator),\n
$$\text{FS}^u_{n}(r_i) := \frac{1}{|G|^n} \sum_{g \in G} \chi_{r_i}(g) u(g_1) \cdots g_n u(g_n),$$  \hspace{1cm} (2.12)\n
where \(\chi_{r_i}\) denotes the character and\n
$$n = \begin{cases} \frac{\text{ord}(u)}{2} & \text{if ord}(u) \text{ is even}, \\ \text{ord}(u) & \text{if ord}(u) \text{ is odd}. \end{cases}$$  \hspace{1cm} (2.13)\n
It has the crucial property\n
$$\text{FS}^u_{n}(r_i) = \pm 1 \quad \forall i \quad \Leftrightarrow \quad u \text{ is class–inverting}.$$  \hspace{1cm} (2.14)\n
So one has to scan over all candidate automorphisms u to determine whether or not one of them is a BDA, a task that can be automatized.

\(^2\)However, the example used there, \(T\), turns out not to be of the CP violating type.

\(^3\)Recall that the Frobenius–Schur indicator \(\text{FS}(r)\) allows one to distinguish between real, pseudo–real and complex representations, for which \(\text{FS}(r)\) takes the values 1, –1 and 0, respectively [11].

Even though the steps in Fig. 2 may, at first sight, appear a bit cumbersome, one should remember that they allow us to uniquely determine, in an automatized way, whether or not a symmetry has a basis in which all CG’s are real, or, if a symmetry clashes with CP. Of course, this analysis is independent of bases, as it should be.

2.7. Three types of groups

Given these tools, one can distinguish between three types of groups [4]:

Case I: for all involutory automorphisms \(u\), i.e., automorphisms that square to unity, of the flavor group there is at least one representation \(r_i\) for which \(\text{FS}_{u_i}(r_i) = 0\). Such discrete symmetries clash with CP.

Case II: there exists an involutory automorphism \(u\) for which the \(\text{FS}_{u_i}\)’s for all representations are non-zero. Then there are two sub–cases:

Case II A: all \(\text{FS}_{u_i}\)’s are +1 for one of those \(u\)’s. In this case, there exists a basis with real Clebsch–Gordan coefficients. The BDA is then the automorphism that corresponds to the physical CP transformation.

Case II B: some of the \(\text{FS}_{u_i}\)’s are –1 for all candidate \(u\)’s. That means that there exists no BDA, and, as a consequence, one cannot find a basis in which all CG’s are real. Nevertheless, any of the \(u\)’s can be used to define a physical CP transformation.

The distinction between the groups is illustrated in Fig. 3.

3. CP VIOLATION WITH AN UNBROKEN CP TRANSFORMATION

Having seen that there are finite groups that do not admit a physical CP transformation, one may wonder about the following question: if one obtains this finite group from a continuous one by spontaneous breaking, at which stage does CP violation arise? That is, take an SU(N) gauge symmetry, impose CP, and break it down to a type–I subgroup. The obvious options how CP violation may come about include

1. CP gets broken by the VEV that breaks SU(N) to G, and
2. the resulting setting always has additional symmetries and does not violate CP.

Rather surprisingly, none of these are the true answer. As demonstrated explicitly in an example in which an SU(3) symmetry gets broken to \(T_7 = Z_3 \times Z_7\), the outer automorphism of SU(3) merges into the outer automorphism of \(T_7\) which however does not entail CP conservation [13].

This leads to a novel way to address the strong CP problem. Start with a theory based on SU(3)\(_C\) × SU(3)\(_F\) (and of course the other gauge symmetries of the Standard Model). Now impose CP, which implies that the coefficient \(\theta\) of the QCD \(C_{\mu \nu} \tilde{G}^\mu \tilde{G}^\nu\) term vanishes. Next, break the continuous flavor symmetry down to a type I flavor symmetry. Then \(\theta\) still vanishes, but CP is violated in the flavor sector. This is required to solve the strong CP problem of the Standard Model. An explicit example will be discussed elsewhere.
4. \(\mathcal{CP}\) VIOLATION FROM STRINGS

Of course, there are alternatives to embedding the discrete flavor symmetry \(G\) into a continuous gauge symmetry (in four spacetime dimensions). In fact, anomaly considerations seem to disfavor this possibility: an \(\text{SU}(3)_F\) symmetry with the families transforming as 3–plets has un-cancelled anomalies (see e.g. [14]). On the other hand, non–Abelian discrete flavor symmetries may originate from extra dimensions [15]. In particular, orbifold compactifications of the heterotic string lead to various flavor groups [16, 17]. These symmetries originate from gauge symmetries in higher dimensions [18, 19], as they should [20]. As it turns out, already the very first 3–generation orbifold model [21] has a \(\Delta(54)\) flavor symmetry [16], which is according to the classification [4] type I and thus \(\mathcal{CP}\)–violating. Therefore, \(\mathcal{CP}\) is violated in such models [22].

When establishing explicitly that \(\mathcal{CP}\) is violated, it was noticed that at the massless level only 1- and 3–dimensional representations of \(\Delta(54)\) occur. There exist outer automorphisms of \(\Delta(54)\) which map all these representations on their conjugates. However, this is no longer the case when one includes the massive states. In particular, the winding strings (see Figure 4) give rise to \(\Delta(54)\) doublets.

The presence of these doublets leads to \(\mathcal{CP}\) violation [22]. This can be made explicit by finding a basis–invariant contraction (see [23]) that has a nontrivial phase. Of course, at this level the flavor symmetry is unbroken, and there is no direct connection between the phase of the contraction presented in [22] and the \(\mathcal{CP}\) violation in the CKM matrix or baryogenesis. One would have to study, how, in explicit models (e.g. [24]) in which the flavor symmetry gets broken and potentially realistic mass matrices arise, this \(\mathcal{CP}\) violation from strings manifests itself in the low–energy effective theory. This has not yet been carried out. Nevertheless, it is clear that if \(\mathcal{CP}\) is broken at the orbifold point, it won’t un-break by moving away from it by e.g., giving the flavons VEVs.

More recently, an additional amusing observation has been made [25]. In orbifold compactifications (without the so–called Wilson lines [26]), the flavor symmetry \(G\) is simply the outer automorphism group of the space group \(S\). In a bit more detail, the states of an orbifold correspond to conjugacy classes of the
space group, and can be represented by space group elements \((\theta^k, n_\alpha e_\alpha)\), where \(\theta^k\) stands for a discrete rotation and \(n_\alpha e_\alpha\) an element of the underlying torus lattice. These conjugacy classes form multiplets under the outer automorphism group of the space group, thus \(G = \text{out}(S)\). This leads to the picture of “out of out”,
\[
\mathcal{CP} \in \text{out}(G) = \text{out}(\text{out}(S))
\]  
(4.1)

Let us also mention that other orbifold geometries come with different flavor symmetries. The probably simplest option is a \(Z_2\) orbifold plane, which leads to a \(D_4\) family symmetry [27, 16]. \(D_4\) is a type II group, meaning that here one cannot immediately conclude that \(\mathcal{CP}\) is violated. On the other hand, it entails a \(Z^A_2\) symmetry, which solves several shortcomings of the supersymmetric Standard Model at once [28, 29, 30, 31]. In particular, it solves the \(\mu\) problem and explains the longevity of the proton and the stability of the LSP. All these examples illustrate the impact of properties of compact dimensions on particle phenomenology.

Arguably, it is rather amusing that \(\mathcal{CP}\) violation can be tied to the presence of states that are required anyway for completing the models in the ultraviolet. One may thus say that, at least in these models, consistency in the ultraviolet requires \(\mathcal{CP}\) to be violated.

5. SUMMARY

\(\mathcal{CP}\) violation may originate from group theory. We have reviewed the observation that there are certain finite groups that clash with \(\mathcal{CP}\) in the sense that, if these groups are realized as (flavor) symmetries, \(\mathcal{CP}\) is violated. To the best of our knowledge, this is a situation that is not too ubiquitous in theory space. What usually happens is that an extra symmetry results from imposing a symmetry. Here, the opposite happens: \(\mathcal{CP}\) can get broken because another (flavor) symmetry is imposed or emerges.

These \(\mathcal{CP}\)-breaking symmetries emerge from explicit string models. Even the earliest 3-generation string models in the literature have a \(\mathcal{CP}\) violating discrete symmetry. In the string models, all symmetries have a clear geometric interpretation, which is why it is fair to say that the origin of \(\mathcal{CP}\) violation described in this review deserves to be called “geometric”.

Acknowledgements

We are indebted to Dillon Berger for valuable comments on this review. We would like to thank Ernest Ma for pushing us to write this up, and the Valencia group for inviting us to their beautiful city. The work of M.C.C. was supported by, in part, by the National Science Foundation under Grant No. PHY-1620638. The work of M.R. is supported by NSF Grant No. PHY-1719438.

References