Proof. This is quite a cute argument: By assumption, we have

$$\frac{1}{|G|}\sum_{g\in G}\chi_\varrho(g^2)=1$$

for all $\varrho\in\widehat{G}$. Multiplying by $\dim(\varrho)$ and then summing over all ϱ, we obtain

$$\begin{aligned}\sum_{\varrho\in\widehat{G}}\dim(\varrho)&=\frac{1}{|G|}\sum_{\varrho\in\widehat{G}}\sum_{g\in G}\chi_\varrho(g^2)\dim(\varrho)\\&=\frac{1}{|G|}\sum_{g\in G}\sum_{\varrho\in\widehat{G}}\chi_\varrho(g^2)\chi_\varrho(1).\end{aligned}$$

By the second orthogonality relation (4.28), the inner sum vanishes unless g^2 is conjugate to 1, i.e., unless $g^2=1$, and in that case it is equal to $|G|$. Thus we get

$$\sum_{\varrho\in\widehat{G}}\dim(\varrho)=|\{g\in G\mid g^2=1\}|,$$

as claimed. □

The problem of evaluating this sum was mentioned in Remark 4.2.6.

Exercise 6.2.7. Consider the examples of finite groups for which we computed the full character table (in Section 4.6.2, 4.6.3 and 4.6.4), and for each of them determine the Frobenius–Schur indicators (in particular determine which are self-dual). (*Hint*: For $\mathrm{GL}_2(\mathbf{F}_p)$, one can use Exercise 4.6.16 to find rather easily the self-dual representations.)

6.3. The Larsen alternative

Our next application has some common features with the Frobenius–Schur theory, but it is a much more recent development which is really a fact about compact, infinite, Lie groups. The results are due to M. Larsen [**42**, §3] and have been extensively developed by N. Katz (for instance in [**32**]).

Their basic motivation can be described as follows. A compact group $G\subset \mathrm{U}_n(\mathbf{C})$ is given, by some means or other, and the question that arises is to identify it, in particular, to prove that it is big in some sense. Here, "big" has roughly the following meaning: either one would like to prove that $G\supset \mathrm{SU}_n(\mathbf{C})$ or one knows—again, one way or another—that G preserves either a symmetric or alternating non-degenerate bilinear form, and the goal is to prove that G contains either the corresponding (real) special orthogonal group or the unitary symplectic group. For this, Larsen found a beautiful numerical criterion. We present it here as an interesting and relatively elementary fact about representations of compact groups. It might not be clear whether this is actually applicable in practice, but we will describe quickly in a later remark how the problem appears in concrete applications.

The invariant introduced by Larsen is the following:

Definition 6.3.1 (Fourth moment of a representation)**.** Let G be a compact subgroup of $\mathrm{U}_n(\mathbf{C})$ for some $n \geqslant 1$ with probability Haar measure μ. The *fourth moment* of G is defined by

$$\mathrm{M}_4(G) = \int_G |\operatorname{Tr}(g)|^4 d\mu(g). \tag{6.8}$$

More generally, given a finite-dimensional representation ϱ of G, the *fourth moment of* ϱ is defined by

$$\mathrm{M}_4(\varrho) = \int_G |\chi_\varrho(g)|^4 d\mu(g).$$

Thus $\mathrm{M}_4(G)$ is the fourth moment of the *tautological* (faithful) representation $\varrho : G \hookrightarrow \mathrm{U}_n(\mathbf{C})$.

A priori, this might be an arbitrary non-negative real number. However, as in the case of the Frobenius–Schur indicator (6.1), it is in fact an integer, and certain of its values carry important meaning. More precisely, we have the following rather remarkable result of Larsen.

Theorem 6.3.2 (Larsen alternative for unitary groups)**.** *Let $n \geqslant 2$, and let G be a compact subgroup of $\mathrm{SU}_n(\mathbf{C})$. If the fourth moment $\mathrm{M}_4(G)$ is equal to 2, then either G is* finite *or $G = \mathrm{SU}_n(\mathbf{C})$. In particular, if G is connected, we have $G = \mathrm{SU}_n(\mathbf{C})$.*

The proof is a very nice application of basic character theory and representation theory, together with some facts of Lie theory. The first step, which we take backwards in comparison with Section 6.2, is to interpret the fourth moment in purely algebraic terms.

Lemma 6.3.3. *Let G be a compact group, and let*

$$\varrho : G \longrightarrow \mathrm{GL}(E)$$

be a finite-dimensional representation of G.

(1) *We have*

$$\mathrm{M}_4(\varrho) = \dim(\operatorname{End}(\varrho) \otimes \operatorname{End}(\varrho))^G = \dim(\operatorname{End}(\varrho \otimes \check{\varrho}))^G. \tag{6.9}$$

(2) *Let π be any of the representations of G on $\varrho \otimes \varrho$, $\varrho \otimes \check{\varrho}$ or $\operatorname{End}(\varrho)$. If we have a decomposition*

$$\pi \simeq \bigoplus_i n_i \varrho_i, \qquad n_i \geqslant 0,$$

into G-stable subspaces where the subrepresentations ϱ_i are not necessarily irreducible, then we have

$$\mathrm{M}_4(\varrho) \geqslant \sum_i n_i^2,$$

with equality if and only if the ϱ_i are pairwise distinct irreducible representations.

(3) *If $G \subset H$ are compact subgroups of $\mathrm{U}_n(\mathbf{C})$, then we have*

$$\mathrm{M}_4(H) \leqslant \mathrm{M}_4(G). \tag{6.10}$$

Proof. Note that the fourth moment is an inner product

$$\mathrm{M}_4(\varrho) = \langle |\chi_\varrho|^4, 1\rangle.$$

By the formalism of characters, the function $|\chi_\varrho|^4$ is the character of the representation

$$\tau = \varrho \otimes \varrho \otimes \check{\varrho} \otimes \check{\varrho},$$

so that $\mathrm{M}_4(\varrho)$ is the dimension of the invariant space τ^G. But using the associativity of the tensor product and the relations

$$(\varrho_1 \otimes \varrho_2)\check{} = \check{\varrho}_1 \otimes \check{\varrho}_2, \qquad \check{\check{\varrho}} = \varrho,$$

we can arrange the tensor product τ in two ways: either

$$\tau = (\varrho \otimes \varrho) \otimes (\varrho \otimes \varrho)\check{} \simeq \mathrm{End}(\varrho \otimes \varrho),$$

which gives

$$\mathrm{M}_4(\varrho) = \dim(\mathrm{End}(\varrho \otimes \varrho))^G,$$

or

$$\tau = (\varrho \otimes \check{\varrho}) \otimes (\varrho \otimes \check{\varrho})\check{} \simeq \mathrm{End}(\varrho \otimes \check{\varrho}) = \mathrm{End}(\mathrm{End}(\varrho)),$$

and therefore

$$\mathrm{M}_4(\varrho) = \dim(\mathrm{End}(\varrho \otimes \check{\varrho}))^G = \dim \mathrm{End}(\mathrm{End}(\varrho))^G.$$

This proves (1), and (2) is a general fact about $\dim \mathrm{End}(\pi)^G$ for any representation π. We have

$$\langle \mathrm{End}(\pi), 1\rangle = \sum_{i,j} n_i n_j \langle \varrho_i, \varrho_j\rangle$$

by linearity. Each term is a non-negative integer, and hence

$$\langle \mathrm{End}(\pi), 1\rangle \geqslant \sum_i n_i^2 \langle \varrho_i, \varrho_i\rangle \geqslant \sum_i n_i^2,$$

by keeping only the diagonal terms $i = j$. If there is equality, we see that we must have $\langle \varrho_i, \varrho_j\rangle = \delta(i,j)$, which means that the ϱ_i are irreducible (taking $i = j$) and distinct (for $i \neq j$).

Finally the inequality (6.10), though not at all obvious from the definition (6.8), is clear from (1): if $G \subset H$, then, for any representation of H, the space of G-invariants contains the space of H-invariants. $\square$

Remark 6.3.4. The reader may have noted that the two quantities

$$\dim(\mathrm{End}(\varrho) \otimes \mathrm{End}(\varrho))^G, \qquad \dim \mathrm{End}(\varrho \otimes \check{\varrho})^G$$

in (6.9) make sense for any (finite-dimensional) representation of any group, and the algebraic argument with associativity of the tensor product used in the proof of the lemma shows that they are equal in this generality. One may therefore *define* an abstract "fourth moment" of a representation using either of them. It is natural to ask, if G is not compact, what is the meaning of this abstract $\mathrm{M}_4(\varrho)$, and in particular (in view of the Larsen alternative) to ask what the equality $\mathrm{M}_4(\varrho) = 2$ means in general. We will discuss this in Section 7.1.

The proof will use a little bit of differential geometry. Readers who are unfamiliar with the notion of manifolds and with the basic definition of Lie groups can simply skim (or skip) the proof.

Proof of the Larsen alternative. To study $\mathrm{M}_4(G)$, we use part (2) of Lemma 6.2.3 for the representation of G on the linear space $\mathrm{End}(\mathbf{C}^n)$, i.e., on $\mathrm{End}(\varrho)$, where

$$\varrho : G \hookrightarrow \mathrm{U}_n(\mathbf{C})$$

is the representation defining G as a subgroup of $\mathrm{U}_n(\mathbf{C})$.

We recall that this representation is the *conjugation* action, i.e., that

$$g \cdot A = gAg^{-1}$$

for $g \in G$ and $A \in E = \mathrm{End}(\mathbf{C}^n)$ (it is the restriction of the corresponding action for $\mathrm{SU}_n(\mathbf{C})$, or indeed for $\mathrm{GL}_n(\mathbf{C})$). There is, as usual, a canonical invariant subspace of dimension one, namely $\mathbf{C}\mathrm{Id} \subset E$. Moreover, a stable (orthogonal) complement is

$$E_0 = \{A \in E \mid \mathrm{Tr}(A) = 0\},$$

the space of endomorphisms of trace 0. Hence we have a first decomposition into subrepresentations

$$E = \mathbf{C}\mathrm{Id} \oplus E_0. \tag{6.11}$$

If only for dimension reasons, the two components are non-isomorphic; therefore, by the previous lemma, we get automatically

$$\mathrm{M}_4(G) \geqslant 1^2 + 1^2 = 2.$$

We deduce first from this that $\mathrm{M}_4(\mathrm{SU}_n(\mathbf{C})) = 2$; indeed, the lemma shows that this simply means that the decomposition (6.11) is a decomposition into irreducible representations of $\mathrm{SU}_n(\mathbf{C})$, which we know is true for the first component because it is one dimensional, and for the second by Exercise 2.7.14.

By the same token, we see that if $G \subset \mathrm{SU}_n(\mathbf{C})$, we can only have $\mathrm{M}_4(G) = 2$ if E_0 is also irreducible as a representation of G. We assume that this is the case and will deduce that G is either finite or equal to $\mathrm{SU}_n(\mathbf{C})$.

To do this, we must appeal to the fact that G is a Lie group, and in fact that it is a smooth manifold.[3] Thus we may consider the tangent space of G at the identity element, which is its Lie algebra,[4] denoted $\mathrm{Lie}(G)$. This is a *real* vector space, of dimension equal to the dimension of G as a manifold. The point is that G acts linearly on $\mathrm{Lie}(G)$, by means of the so-called *adjoint* representation,[5] which is obtained by differentiating at the identity the conjugation action of G on itself: denoting by $I(g)$ the inner automorphism that maps x to gxg^{-1}, the adjoint representation is given by

$$\mathrm{Ad}\begin{cases} G & \longrightarrow & \mathrm{GL}(\mathrm{Lie}(G)) \\ g & \mapsto & dI(g)_e, \end{cases}$$

where $dI(g)_e$ denotes the tangent map of the diffeomorphism $I(g)$ at the identity element e of G. This is a well-defined linear map on $\mathrm{Lie}(G)$ (since $I(g)(e) = e$ for each g), and it is a representation, by the chain rule, because $I(gh) = I(g)I(h)$ and $I(g^{-1}) = I(g)^{-1}$.

This representation is a *real* representation[6] of G, since $\mathrm{Lie}(G)$ is a real vector space. Most crucial for us, it has the following property, which is almost immediate: if $G \subset H$ with H also a compact Lie group, then $\mathrm{Lie}(G) \subset \mathrm{Lie}(H)$, and the adjoint representation of G is the restriction of the adjoint representation of H. Applied to $G \subset \mathrm{SU}_n(\mathbf{C})$, it follows that $\mathrm{Lie}(G)$ is a subrepresentation of $\mathrm{Lie}(\mathrm{SU}_n(\mathbf{C}))$.

This is the source of the desired subrepresentation of E_0. We will check below the following facts:

- The Lie algebra L_n of $\mathrm{SU}_n(\mathbf{C})$ is a real subspace of E_0, such that $L_n \oplus iL_n = L_n \otimes \mathbf{C} = E_0$;

[3] It does not suffice here to know that G is a topological manifold.

[4] Although we will in fact not need the structure of Lie algebra (see Section 3.2) that exists on this space.

[5] It may be confusing at first that there there exist an adjoint representation for a Lie group and one for a Lie algebra (see Example 3.2.3) and that neither has much to do with the adjoint of a linear map.

[6] Not to be confused with representations of *real type*, which were mentioned briefly in Definition 6.2.2.

- The adjoint representation of $\mathrm{SU}_n(\mathbf{C})$ on L_n is a real subrepresentation of E_0, i.e., on $L_n \subset E_0$, the adjoint representation is given by $g \cdot A = gAg^{-1}$ for $A \in L_n \subset E_0$.

If we assume these, we can conclude as follows. For our compact subgroup $G \subset \mathrm{SU}_n(\mathbf{C})$, we have the subrepresentation

$$\mathrm{Lie}(G) \otimes \mathbf{C} \subset L_n \otimes \mathbf{C} = E_0$$

of the G-action on E_0. Since we are assuming that the latter is irreducible, this means that either $\mathrm{Lie}(G)$ is 0, in which case G is finite, or that $\mathrm{Lie}(G)$ is equal to L_n, in which case, by Lie theory, we have $G = \mathrm{SU}_n(\mathbf{C})$. Hence the Larsen alternative is proved.

Now we explain the facts mentioned above—these are quite standard and the reader may well have already encountered them. To begin with, the special unitary group is defined by the conditions

$$\det(g) = 1, \qquad gg^* = 1,$$

in $\mathrm{GL}_n(\mathbf{C})$. The tangent space at 1 is obtained by considering the linearized forms of these equations, viewed as applying to matrices A in $\mathrm{M}_n(\mathbf{C})$, which form the tangent space at 1 of $\mathrm{GL}_n(\mathbf{C})$. The first equation becomes $\mathrm{Tr}(A) = 0$, which means $A \in E_0$, and the second becomes

$$A + A^* = 0,$$

i.e., A is skew-hermitian, so

$$L_n = \{A \in \mathrm{M}_n(\mathbf{C}) \mid A = -A^*, \text{ and } \mathrm{Tr}(A) = 0\} \subset E_0. \tag{6.12}$$

(Note that since the adjoint operation $A \mapsto A^*$ is not complex linear, this is indeed only a real vector space.)

We can easily check explicitly that $E_0 = L_n \otimes \mathbf{C}$: for $A \in E_0$, we write

$$A = \frac{A + A^*}{2} + \frac{A - A^*}{2} = iB + C \quad \text{(say)}.$$

Then $C^* = -C$, so C is skew-hermitian, and $B = (A + A^*)/(2i)$ has also $B^* = -(A^* + A)/(2i) = -B$, so that B is skew-hermitian. Since $\mathrm{Tr}(B) = \mathrm{Re}(\mathrm{Tr}(A))$ and $\mathrm{Tr}(C) = i\,\mathrm{Im}(\mathrm{Tr}(A))$, we deduce $\mathrm{Tr}(B) = \mathrm{Tr}(C) = 0$, so that we have found a decomposition of A as $C + iB$ with C, B both in L_n. This decomposition is unique, because $L_n \cap iL_n = 0$ (in E_0): any matrix in the intersection is both hermitian and skew-hermitian. So this proves the first claim.

The second one is not too surprising since the adjoint representation is defined using conjugation. To be precise, let $A \in L_n$ be a tangent vector. Then elementary differential geometry tells us that $\mathrm{Ad}(g)(A)$ can be computed as

$$\frac{d}{dt}I(g)(x_t)\Big|_{t=0},$$

where $x_t \in \mathrm{SU}_n(\mathbf{C})$ defines any smooth curve with tangent vector A at $t = 0$. As usual, one takes $x_t = \exp(tA)$, where the exponential is that of matrices; then we have

$$I(g)x_t = g\exp(tA)g^{-1} = \exp(tgAg^{-1})$$

(e.g., using the Taylor series expansion), and the derivative at $t = 0$ gives $\mathrm{Ad}(g)A = gAg^{-1}$, as desired. □

Example 6.3.5 (Finite groups with $\mathrm{M}_4 = 2$)**.** As observed by Katz [**32**, 1.6.1], there do exist finite groups $G \subset \mathrm{SU}_n(\mathbf{C})$, for some $n \geqslant 2$, for which $\mathrm{M}_4(G) = 2$. For instance, let $G = \mathrm{PSL}_2(\mathbf{F}_7)$. It follows from the character table of $\mathrm{SL}_2(\mathbf{F}_7)$ that G has two distinct irreducible representations π_1 and π_2 of dimension $3 = (7-1)/2$. Unitarized, either of these gives a homomorphism

$$G \longrightarrow \mathrm{U}_3(\mathbf{C}).$$

Since G is a simple group, this is necessarily a faithful representation, and (for the same reason) the composite $G \hookrightarrow \mathrm{U}_3(\mathbf{C}) \xrightarrow{\det} \mathbf{C}^\times$, which cannot be injective, is trivial. Thus the image of either of these representations is a finite subgroup of $\mathrm{U}_3(\mathbf{C})$, and one can check that these have a fourth moment equal to 2.

In addition to the case of the unitary group considered above, there are criteria for orthogonal and symplectic groups. Let $n \geqslant 4$ be an even integer, and let $G \subset \mathrm{U}_n(\mathbf{C})$ be a connected compact group which is contained in the subgroup $\mathrm{USp}_n(\mathbf{C}) = \mathrm{Sp}_n(\mathbf{C}) \cap \mathrm{U}_n(\mathbf{C})$ of unitary matrices that leave invariant a non-degenerate alternating bilinear form. Then Larsen showed that $G = \mathrm{USp}_n(\mathbf{C})$ if and only if $\mathrm{M}_4(G) = 3$. Similarly, if $n \geqslant 2$ and $G \subset \mathrm{U}_n(\mathbf{C})$ is a compact connected group contained in the compact group $\mathrm{O}_n(\mathbf{R}) = \mathrm{U}_n(\mathbf{C}) \cap \mathrm{O}_n(\mathbf{C})$ of unitary matrices leaving invariant a non-degenerate symmetric bilinear form, we have $G \supset \mathrm{SO}_n(\mathbf{R})$ if and only if $\mathrm{M}_4(G) = 3$. One cannot, however, distinguish between $\mathrm{SO}_n(\mathbf{R})$ and $\mathrm{O}_n(\mathbf{R})$ using the fourth moment.

The point in these results is that $\mathrm{USp}_n(\mathbf{C})$ and $\mathrm{O}_n(\mathbf{R})$ are maximal compact subgroups, respectively, of $\mathrm{Sp}_n(\mathbf{C})$ and $\mathrm{O}_n(\mathbf{C})$.

We sketch the argument for the symplectic case: denoting $V = \mathbf{C}^n$, one has a decomposition

$$V^{\otimes 2} = \mathrm{Sym}^2(V) \oplus \mathbf{C} \oplus V_1$$

as the $\mathrm{USp}_n(\mathbf{C})$-representation, where the trivial one-dimensional component $\mathbf{C}$ corresponds to the (dual of the) invariant alternating form on V, and $V_1 \neq 0$ because $\dim V \geqslant 4$. As representations of $\mathrm{USp}_n(\mathbf{C})$, the three pieces are known to be irreducible, so that $\mathrm{M}_4(\mathrm{USp}_n(\mathbf{C})) = 3$, and by Lemma 6.3.3(1), we have $\mathrm{M}_4(G) = 3$ if and only if all three representations are G-irreducible. It also turns out that $\mathrm{Sym}^2(V)$ is isomorphic to the adjoint representation of $\mathrm{USp}_n(\mathbf{C})$ on its Lie algebra, and hence it contains as a G-invariant subspace the Lie algebra of G itself. Therefore, irreducibility of $\mathrm{Sym}^2(V)$ implies that $G = \mathrm{USp}_n(\mathbf{C})$ since both are connected with the same Lie algebra.

Finally, we address the problem of applications of the Larsen alternative. We explain here, with a specific example, some of the situations where results like this are very valuable tools. As already hinted, sometimes theory gives the existence of some group which carries information concerning objects of interest. A very good example, though it is not directly relevant to the Larsen alternative, is the Galois group of the splitting field of a polynomial. This is a finite group, which is (usually) defined rather abstractly, so that if one knows the coefficients of the polynomial, it is not easy at all to determine the Galois group. In fact, often the only obvious information is that it is isomorphic to a subgroup of $\mathfrak{S}_n$, where n is the degree of the polynomial. For instance, can you guess the Galois group of the splitting field of

$$X^8 - 4X^7 + 8X^6 - 11X^5 + 12X^4 - 10X^3 + 6X^2 - 3X + 2$$

over $\mathbf{Q}$?

Now for the example, which is based on deep work and ideas of P. Deligne [**15**] and of N. Katz [**33**], and for which we assume some familiarity with the basic theory of finite fields (again, [**54**, Chapter 1] is an excellent reference).

Let $q = p^d$ be a prime power, and let $\mathbf{F}_q$ be a finite field with q elements. The map

$$\mathrm{Tr}_{\mathbf{F}_q/\mathbf{F}_d} \left\{ \begin{array}{ccc} \mathbf{F}_q & \longrightarrow & \mathbf{F}_p \\ x & \mapsto & x + x^p + \cdots + x^{p^{d-1}} \end{array} \right.$$

is well defined and is a homomorphism of additive groups (the element $y = \mathrm{Tr}_{\mathbf{F}_q/\mathbf{F}_p}$ is in $\mathbf{F}_p$ because it satisfies $y^p = y$). We denote by ψ the non-trivial additive character of the additive group of $\mathbf{F}_p$ defined by

$$\psi(x) = e\Big(\frac{x}{p}\Big)$$

for $x \in \mathbf{F}_p$ (see Remark 4.5.3). The composition

$$\psi_{\mathbf{F}_q/\mathbf{F}_d} = \psi \circ \mathrm{Tr}_{\mathbf{F}_q/\mathbf{F}_p}$$

is then a character of the additive group of $\mathbf{F}_q$, and since one can show that the trace map $\mathrm{Tr}_{\mathbf{F}_q/\mathbf{F}_p}$ is surjective, it follows that $\psi_{\mathbf{F}_q/\mathbf{F}_p}$ is also non-trivial.

Note that if d is not divisible by p, the surjectivity follows easily from the fact that $\mathrm{Tr}_{\mathbf{F}_q/\mathbf{F}_p}(1) = d$ is then non-zero in $\mathbf{F}_p$.

Now, for any element $a \in \mathbf{F}_q$, one defines

$$S(a;q) = \sum_{x \in \mathbf{F}_q^{\times}} \psi_{\mathbf{F}_q/\mathbf{F}_p}(ax + x^{-1}), \tag{6.13}$$

a sum which is called a *Kloosterman sum*.

These sums are apparently just complex numbers, but they turn out to be related to some compact Lie groups. Indeed, it follows from the work of A. Weil that for every $a \in \mathbf{F}_q$ there exists a well-defined *conjugacy class* $\theta(a;q)$ in the unitary group $\mathrm{U}_2(\mathbf{C})$ such that

$$\mathrm{Tr}\,\theta(a;q) = -\frac{S(a;q)}{\sqrt{q}}, \tag{6.14}$$

where the trace on the left-hand side is just the ordinary trace of matrices, which can be applied to conjugacy classes, since the trace is invariant under conjugation. In particular, note that this formula immediately implies the bound

$$|S(a;q)| \leqslant 2\sqrt{q},$$

for $a \in \mathbf{F}_q^{\times}$, which is a deep theorem of A. Weil. In fact, it is a special case of the *Riemann hypothesis for curves over finite fields*; we refer, e.g., to [**30**, Th. 11.11] for one of the simplest proofs.

The connection with the Larsen alternative arises from the following fact, which is a special case of a famous theorem of Deligne (Deligne's equidistribution theorem) applied to the so-called Kloosterman sheaves. There exists a compact subgroup $K \subset \mathrm{U}_2(\mathbf{C})$, depending a priori on p, such that, first, all $\theta(a;q)$ are in fact naturally conjugacy classes of K, and second, they become *equidistributed* among conjugacy classes of K, in the sense that for any continuous class function $f : K \longrightarrow \mathbf{C}$, we have

$$\int_K f(x)d\mu(x) = \lim_{q\to+\infty} \frac{1}{q-1} \sum_{a \in \mathbf{F}_q^{\times}} f(\theta(a;q)), \tag{6.15}$$

where μ is the probability Haar measure on K, and q tends to infinity through powers of p. See the discussion in [**33**, Ch. 3].

Thus, if one succeeds in determining what the group K is—something which, just as was the case for Galois group, is by no means clear by just looking at (6.13)!—one can answer many questions about the asymptotic distribution of Kloosterman sums, something which is of great interest in number theory.

Now it is clear why the Larsen alternative is useful: applying first (6.15) with $f(x) = |\operatorname{Tr}(x)|^4$ and then (6.14), we get the alternative formula

$$\begin{aligned} \mathrm{M}_4(K) &= \lim_{q\to+\infty} \frac{1}{q-1} \sum_{a\in\mathbf{F}_q^\times} |\operatorname{Tr}\theta(a;q)|^4 \\ &= \lim_{q\to+\infty} \frac{1}{q^2(q-1)} \sum_{a\in\mathbf{F}_q^\times} \Big| \sum_{x\in\mathbf{F}_q^\times} \psi_{\mathbf{F}_q/\mathbf{F}_p}(ax+x^{-1}) \Big|^4 \end{aligned}$$

for the fourth moment of K, which involves the given concrete data defining the problem. We may have a chance to evaluate this....

As it turns out, one can evaluate directly this limit in this case (this is a relatively elementary computation, see Exercise 6.3.6 below) and see that it exists and is equal to 2. In other words, the compact group K satisfies $\mathrm{M}_4(K) = 2$. Hence the Larsen alternative shows that either K is finite, or $K \supset \mathrm{SU}_2(\mathbf{C})$. In fact, one can analyze the situation further, and show that K is equal to the special unitary group $\mathrm{SU}_2(\mathbf{C})$. It follows then from (6.15), for instance, that

$$\lim_{q\to+\infty} \frac{1}{q^{k+1/2}(q-1)} \sum_{a\in\mathbf{F}_q^\times} \Big| \sum_{x\in\mathbf{F}_q^\times} \psi_{\mathbf{F}_q/\mathbf{F}_p}(ax+x^{-1}) \Big|^{2k+1} = 0$$

for any integer $k \geqslant 0$ and

$$\lim_{q\to+\infty} \frac{1}{q^{k}(q-1)} \sum_{a\in\mathbf{F}_q^\times} \Big| \sum_{x\in\mathbf{F}_q^\times} \psi_{\mathbf{F}_q/\mathbf{F}_p}(ax+x^{-1}) \Big|^{2k} = \frac{1}{k+1}\binom{2k}{k}$$

for any integer $k \geqslant 1$ (see Exercise 6.3.7(1) below).

This may seem too good to be true, but we remind the reader however that the *existence* of this group K lies extremely deep: the equidistribution formula (6.15) cannot be proven without first knowing its existence, and only later can one attempt to determine the group K.

In the works of Katz, many other (more general) situations are considered, leading to extremely general and beautiful equidistribution theorems. But even though the statements can be extremely concrete, there is no known elementary proof of the deep connection between Kloosterman sums (or other similar sums) and a compact Lie group!

Exercise 6.3.6 (Fourth moment of Kloosterman sums)**.** For a finite field $\mathbf{F}_q$ with q elements and ψ as above, define

$$T(a,b;q) = \sum_{x\in\mathbf{F}_q^\times} \psi_{\mathbf{F}_q/\mathbf{F}_p}(ax+bx^{-1}).$$

(1) Show that $T(a,b;q) = S(ab;q)$ if a, $b \in \mathbf{F}_q^\times$.

(2) Let

$$N_4(q) = \sum_{a \in \mathbf{F}_q^\times} |S(a;q)|^4.$$

Deduce from (1) that

$$(q-1)N_4(q) = q^2 \mathcal{N}(q) - (q-1)^4 - 2(q-1),$$

where

$$\mathcal{N}(q) = \Big|\Big\{(x_1,x_2,y_1,y_2) \in (\mathbf{F}_q^\times)^4 \mid x_1+x_2 = y_1+y_2 \text{ and } \frac{1}{x_1}+\frac{1}{x_2} = \frac{1}{y_1}+\frac{1}{y_2}\Big\}\Big|.$$

(3) Prove that

$$\mathcal{N}(q) = 3(q-2)(q-1),$$

and deduce that

$$\lim_{q\to+\infty} \frac{N_4(q)}{q-1} = 2.$$

(*Hint*: Use the fact that a pair $(x+y, xy)$ determines x and y up to order.)

Exercise 6.3.7 (Other moments)**.** One can define other types of moments. For instance, given a compact group G (with probability Haar measure μ) and a finite-dimensional unitary representation ϱ of G, the k-th moment of ϱ is defined to be

$$M_k(\varrho) = \int_G \chi_\varrho(g)^k d\mu(g)$$

for an integer $k \geqslant 0$. If $G \subset \mathrm{GL}_n(\mathbf{C})$ is a compact subgroup, we denote by $M_k(G)$ the k-th moment of this inclusion.

It is an elementary consequence of character theory, which is not necessarily clear at first when expressed for a concrete group, that $M_k(\varrho)$ is a non-negative integer, since it is the multiplicity of the trivial representation in the finite-dimensional representation $\varrho^{\otimes k}$ (see (5.12)).

The sequence of moments $(M_k(\varrho))_{a\geqslant 0}$, as $k \geqslant 0$ varies, can be quite interesting:

(1) Take $G = \mathrm{SU}_2(\mathbf{C}) \subset \mathrm{GL}_2(\mathbf{C})$. Show that

$$M_k(G) = 0$$

if k is odd and

$$M_{2k}(G) = \frac{1}{k+1}\binom{2k}{k}$$

for $k \geqslant 0$. Can you prove directly that the right-hand side is an integer?

(2) Compute the first few terms and identify this sequence in the *Online Encyclopedia of Integer Sequences* (`http://oeis.org`). (These are called the *Catalan numbers* and have extremely varied interpretations.)

(3) Let $G = \mathrm{SU}_n(\mathbf{C}) \subset \mathrm{GL}_n(\mathbf{C})$. Show that $M_k(G) \neq 0$ if and only if k is divisible by n. What happens when $G = \mathrm{U}_n(\mathbf{C})$?

Exercise 6.3.8 (Another application of the Larsen alternative)**.** For a prime number p and an element $a \in \mathbf{F}_p^\times$, let

$$S_3(a;p) = \frac{1}{p}\sum\sum_{x,y\in\mathbf{F}_p^\times} e\Bigl(\frac{x+y+a(xy)^{-1}}{p}\Bigr),$$

which is called a a hyper-Kloosterman sum in two variables (the inverse $(xy)^{-1}$ is computed in $\mathbf{F}_p^\times$).

(1) For reasonably large values of p (say $p \leqslant 100000$) and the first few $k \geqslant 0$, compute (using a computer) the *empirical* moments

$$m_{k,p} = \frac{1}{p-1}\sum_{\alpha\in\mathbf{F}_p^\times} S_3(a,p)^k.$$

Discuss the behavior of the result as p grows.

(2) Can you make a guess concerning some analogue of the equidistribution result for Kloosterman sums discussed above? Check in [**33**] whether this guess is correct.

Exercise 6.3.9 (Maximal fourth moment)**.** (1) Let $G \subset \mathrm{U}_n(\mathbf{C})$ be a compact subgroup of $\mathrm{U}_n(\mathbf{C})$, and let $\varrho : G \longrightarrow \mathrm{U}_m(\mathbf{C})$ be an irreducible unitary representation of G. Show that $\mathrm{M}_4(\varrho) \leqslant m^2$. If G is connected, show that equality holds if and only if $m = 1$.

(2) Show that the dihedral group D_4 of order 8 has a two-dimensional irreducible representation ϱ with $\mathrm{M}_4(\varrho) = 4$.

Remark 6.3.10 (From $\mathrm{SU}_2(\mathbf{C})$ to $\mathrm{SO}_3(\mathbf{R})$)**.** The adjoint representation turns out to provide a conceptual explanation of the projection homomorphism

$$p : \mathrm{SU}_2(\mathbf{C}) \longrightarrow \mathrm{SO}_3(\mathbf{R})$$

of Proposition 5.6.9. Indeed, for the compact Lie group $G = \mathrm{SU}_2(\mathbf{C})$, the Lie algebra is a three-dimensional real vector space (by (6.12) $\mathrm{M}_2(\mathbf{C})$ has dimension 8, the skew-hermitian condition implies that the bottom-left coefficient is minus the conjugate of the top-right one, and the diagonal ones are purely imaginary, leaving $8-2-2=4$ dimensions, and the matrices of trace zero form a three-dimensional subspace). In fact,

$$L_2 = \Bigl\{\begin{pmatrix} ia & c+id \\ -c+id & -ia\end{pmatrix} \mid a,\ c,\ d \in \mathbf{R}\Bigr\},$$

so that a matrix representation for the adjoint representation of $\mathrm{SU}_2(\mathbf{C})$ on L_2 is a homomorphism

$$\mathrm{Ad}^{\boldsymbol{m}} : \mathrm{SU}_2(\mathbf{C}) \longrightarrow \mathrm{GL}_3(\mathbf{R}).$$

This *is* the desired projection, in the sense that it has kernel $\{\pm 1\}$, and image conjugate to $\mathrm{SO}_3(\mathbf{R})$ in $\mathrm{GL}_3(\mathbf{R})$ (depending on which basis of the Lie algebra L_2 is used to compute the matrix form of the representation).

In topological terms, the projection p is a non-trivial covering map of $\mathrm{SO}_3(\mathbf{R})$ (since $\mathrm{SU}_2(\mathbf{C})$ is connected). Thus $\mathrm{SO}_3(\mathbf{R})$ is *not simply connected* (in fact, one can show that $\mathrm{SU}_2(\mathbf{C})$ is simply connected, so it is the universal covering of $\mathrm{SO}_3(\mathbf{R})$). There are well-known *physical* demonstrations of this property of the rotation group (due in particular to Dirac); see, e.g., [**6**] for an accessible mathematical account, though seeing movies on the web might be even more enlightening.. . .

6.4. The hydrogen atom

We now come to the discussion of Example 1.2.3, i.e., of the basic invariants of simple quantum-mechanical systems, and in particular of the hydrogen atom.

In order to do this, we summarize briefly the fundamental formalism of (non-relativistic) quantum mechanics, contrasting it with classical Newtonian mechanics, in the simplest situation of a single (point-like) particle evolving in $\mathbf{R}^3$, under the influence of some force (or forces):

- The *state* of the system at a given time t is represented by a unit vector ψ (i.e., with $\|\psi\| = 1$) in some fixed complex Hilbert space H. In contrast, in Newtonian mechanics, the state of the particle is represented by an element $(x, p) \in \mathbf{R}^6$, where x represents the position of the particle and p its momentum $p = mv$, where $v \in \mathbf{R}^3$ is the the velocity at t and m is the mass of the particle.
- Two vectors ψ_1, ψ_2 in H correspond to the same state if and only if there exists $\theta \in \mathbf{R}$ such that $\psi_1 = e^{i\theta}\psi_2$, i.e., if the vectors are proportional.
- An *observable quantity* (or just an *observable*), such as position or momentum, is represented by a linear operator A defined on a dense subspace D_A of H. If A is continuous, it can be defined on all of H, but many interesting observables are not continuous on D_A. Moreover A must be *self-adjoint*, which has the usual meaning when A is continuous on H, and has a more technical definition otherwise (see Exercise 6.4.1). On the other hand, in Newtonian mechanics, an observable quantity is simply a real-valued function $f : P \longrightarrow \mathbf{R}$, where $P \subset \mathbf{R}^6$ is the set of possible states of the system.
- The physical interaction of the system described by the state ψ with an observable A must result, through experiments, in some actual