1

Haar Measure on the Classical Compact Matrix Groups

1.1 The Classical Compact Matrix Groups

The central objects of study in this book are randomly chosen elements of the classical compact matrix groups: the orthogonal group $\mathbb{O}(n)$, the unitary group $\mathbb{U}(n)$, and the symplectic group $\mathbb{Sp}(2n)$. The groups are defined as follows.

Definition

1. An $n \times n$ matrix U over $\mathbb{R}$ is **orthogonal** if
$$UU^T = U^T U = I_n, \tag{1.1}$$
where I_n denotes the $n \times n$ identity matrix, and U^T is the transpose of U. The set of $n \times n$ orthogonal matrices over $\mathbb{R}$ is denoted $\mathbb{O}(n)$.
2. An $n \times n$ matrix U over $\mathbb{C}$ is **unitary** if
$$UU^* = U^* U = I_n, \tag{1.2}$$
where U^* denotes the conjugate transpose of U. The set of $n \times n$ unitary matrices over $\mathbb{C}$ is denoted $\mathbb{U}(n)$.
3. A $2n \times 2n$ matrix U over $\mathbb{C}$ is **symplectic** if $U \in \mathbb{U}(2n)$ and
$$UJU^T = U^T JU = J, \tag{1.3}$$
where
$$J := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}. \tag{1.4}$$
The set of $2n \times 2n$ symplectic matrices over $\mathbb{C}$ is denoted $\mathbb{Sp}(2n)$.

   Alternatively, the symplectic group can be defined as the set of $n \times n$ matrices U with quaternionic entries, such that $UU^* = I_n$, where U^* is the (quaternionic) conjugate transpose: for
$$\mathbb{H} = \{a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} : a, b, c, d \in \mathbb{R}\}$$

the skew-field of quaternions, satisfying the relations

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1,$$

quaternionic conjugation is defined by

$$\overline{a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}.$$

Quaternions can be represented as 2×2 matrices over $\mathbb{C}$: the map

$$a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \longmapsto \begin{bmatrix} a + bi & c + di \\ -c + di & a - bi \end{bmatrix}$$

is an isomorphism of $\mathbb{H}$ onto

$$\left\{ \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} : z, w \in \mathbb{C} \right\}.$$

More generally, if $A, B, C, D \in \mathrm{M}_n(\mathbb{R})$, then the matrix

$$M = A + B\mathbf{i} + C\mathbf{j} + D\mathbf{k} \in \mathrm{M}_n(\mathbb{H})$$

is associated to the matrix

$$M_{\mathbb{C}} = I_2 \otimes A + iQ_2 \otimes B + Q_3 \otimes C + iQ_4 \otimes D,$$

where

$$Q_2 := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad Q_3 := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad Q_4 := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

and $\otimes$ denotes the Kronecker product. Any matrix $M \in \mathrm{M}_{2n}(\mathbb{C})$ of this form has the property that

$$MJ = J\overline{M}$$

for $J = Q_3 \otimes I_n$ as above, and the condition $UU^* = I_n$ for $U \in \mathrm{M}_n(\mathbb{H})$ is equivalent to $U_{\mathbb{C}}U_{\mathbb{C}}^* = I_n$ over $\mathbb{C}$.

We will generally consider the symplectic group in its complex version, as a subgroup of the (complex) unitary group, although certain geometric properties of the group can be more cleanly characterized in the quaternionic form.

Note that it is immediate from the definitions that U is orthogonal if and only if U^T is orthogonal, and U is unitary or symplectic if and only if U^* is.

The algebraic definitions given above are nicely compact but may not make the importance of these groups jump right out; the following lemma gives some indication as to why they play such a central role in many areas of mathematics.

Lemma 1.1 *1. Let M be an $n \times n$ matrix over $\mathbb{R}$ or $\mathbb{C}$. Then M is orthogonal or unitary if and only if the columns of M form an orthonormal basis of $\mathbb{R}^n$, resp. $\mathbb{C}^n$.*

2. For U an $n \times n$ matrix over $\mathbb{R}$, $U \in \mathbb{O}(n)$ if and only if U acts as an isometry on $\mathbb{R}^n$; that is,

$$\langle Uv, Uw\rangle = \langle v, w\rangle$$

for all $v, w \in \mathbb{R}^n$.

3. For U an $n \times n$ matrix over $\mathbb{C}$, $U \in \mathbb{U}(n)$ if and only if U acts as an isometry on $\mathbb{C}^n$:

$$\langle Uv, Uw\rangle = \langle v, w\rangle$$

for all $v, w \in \mathbb{C}^n$.

4. Consider $\mathbb{C}^{2n}$ equipped with the skew-symmetric form

$$\omega(v, w) = v_1 w_{n+1} + \cdots + v_n w_{2n} - v_{n+1} w_1 - \cdots - v_{2n} w_n = \sum_{k,\ell} J_{kl} v_k w_\ell,$$

where

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

as above. For a $2n \times 2n$ matrix U over $\mathbb{C}$, $U \in \mathbb{Sp}(2n)$ if and only if U is an isometry of $\mathbb{C}^{2n}$ which preserves ω:

$$\langle Uv, Uw\rangle = \langle v, w\rangle \quad \text{and} \quad \omega(Uv, Uw) = \omega(v, w)$$

for all $v, w \in \mathbb{C}^{2n}$.

5. If $U \in \mathbb{O}(n)$ or $U \in \mathbb{U}(n)$, then $|\det(U)| = 1$. If $U \in \mathbb{Sp}(2n)$, then $\det(U) = 1$.

Proof Note that the $(i,j)^{th}$ entry of $U^T U$ (if U has real entries) or $U^* U$ (if U has complex or quaternionic entries) is exactly the inner product of the ith and jth columns of U. So $U^T U = I_n$ or $U^* U = I_n$ is exactly the same thing as saying the columns of U form an orthonormal basis of $\mathbb{R}^n$ or $\mathbb{C}^n$.

For $U \in \mathrm{M}_n(\mathbb{R})$, $\langle Uv, Uw\rangle = \langle U^T Uv, w\rangle$, and so $\langle Uv, Uw\rangle = \langle v, w\rangle$ for all v and w if and only if $U^T U = I$. The proofs of parts 3 and 4 are similar. For part 5, on any of the groups,

$$|\det(U)|^2 = \det(U)\overline{\det(U)} = \det(U)\det(U^*) = \det(UU^*) = \det(I_n) = 1.$$

The easiest way to see that if $U \in \mathbb{Sp}(2n)$, then in fact $\det(U) = 1$ is to use the Pfaffian: for a skew-symmetric matrix A, the Pfaffian $\mathrm{pf}(A)$ is defined by a

sum-over-permutations formula along the lines of the determinant, and has the property that for $2n \times 2n$ matrices A and B,

$$\mathrm{pf}(BAB^T) = \det(B)\,\mathrm{pf}(A).$$

Applying this to the defining relation of $\mathbb{Sp}(2n)$,

$$\mathrm{pf}(J) = \mathrm{pf}(UJU^T) = \det(U)\,\mathrm{pf}(J),$$

and so (using the easily verified fact that $\mathrm{pf}(J) \neq 0$), $\det(U) = 1$. □

We sometimes restrict attention to the "special" counterparts of the orthogonal and unitary groups, defined as follows.

Definition The set $\mathbb{SO}(n) \subseteq \mathbb{O}(n)$ of **special orthogonal matrices** is defined by

$$\mathbb{SO}(n) := \{U \in \mathbb{O}(n) : \det(U) = 1\}.$$

The set $\mathbb{SO}^-(n) \subseteq \mathbb{O}(n)$ (the **negative coset**) is defined by

$$\mathbb{SO}^-(n) := \{U \in \mathbb{O}(n) : \det(U) = -1\}.$$

The set $\mathbb{SU}(n) \subseteq \mathbb{U}(n)$ of **special unitary matrices** is defined by

$$\mathbb{SU}(n) := \{U \in \mathbb{U}(n) : \det(U) = 1\}.$$

Since the matrices of the classical compact groups all act as isometries of $\mathbb{C}^n$, all of their eigenvalues lie on the unit circle $\mathbb{S}^1 \subseteq \mathbb{C}$. In the orthogonal and symplectic cases, there are some built-in symmetries:

Exercise 1.2 Show that each matrix in $\mathbb{SO}(2n+1)$ has 1 as an eigenvalue, each matrix in $\mathbb{SO}^-(2n+1)$ has -1 as an eigenvalue, and each matrix in $\mathbb{SO}^-(2n+2)$ has both -1 and 1 as eigenvalues.

The sets $\mathbb{O}(n)$, $\mathbb{U}(n)$, $\mathbb{Sp}(2n)$, $\mathbb{SO}(n)$, and $\mathbb{SU}(n)$ of matrices defined above are *compact Lie groups*; that is, they are groups (with matrix multiplication as the operation), and they are compact manifolds, such that the multiplication and inverse maps are smooth. Moreover, these groups can naturally be viewed as closed submanifolds of Euclidean space: $\mathbb{O}(n)$ and $\mathbb{SO}(n)$ are submanifolds of $\mathbb{R}^{n^2}$; $\mathbb{U}(n)$ and $\mathbb{SU}(n)$ are submanifolds of $\mathbb{C}^{n^2}$; and $\mathbb{Sp}(2n)$ is a submanifold of $\mathbb{C}^{(2n)^2}$. Rather than viewing these matrices as n^2-dimensional vectors, it is more natural to view them as elements of the Euclidean spaces $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) of $n \times n$ matrices over $\mathbb{R}$ (resp. $\mathbb{C}$), where the Euclidean inner products are written as

$$\langle A, B\rangle_{HS} := \mathrm{Tr}(AB^T)$$

for $A, B \in M_n(\mathbb{R})$, and

$$\langle A, B\rangle_{HS} := \mathrm{Tr}(AB^*)$$

for $A, B \in M_n(\mathbb{C})$. These inner products are called the **Hilbert–Schmidt inner products** on matrix space.

The Hilbert–Schmidt inner product induces a norm on matrices; it is sometimes called the Frobenius norm or the Schatten 2-norm, or just the Euclidean norm. This norm is *unitarily invariant*:

$$\|UBV\|_{HS} = \|B\|_{HS}$$

when U and V are unitary (as is easily seen from the definition). This implies in particular that if $U \in \mathbb{O}(n)$ (resp. $\mathbb{U}(n)$), then the map $R_U : M_n(\mathbb{R}) \to M_n(\mathbb{R})$ (resp. $R_U : M_n(\mathbb{C}) \to M_n(\mathbb{C})$) defined by

$$R_U(M) = UM$$

is an isometry on $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) with respect to the Hilbert–Schmidt inner product.

The Hilbert–Schmidt norm is also *submultiplicative*:

$$\|AB\|_{HS} \leq \|A\|_{HS}\|B\|_{HS}.$$

In fact, this is true of all unitarily invariant norms (subject to the normalization $\|E_{11}\| = 1$), but it is particularly easy to see for the Hilbert–Schmidt norm: let B have columns $b_1, \ldots, b_n$; then $\|B\|_{HS}^2 = \sum_{j=1}^n |b_j|^2$, where $|\cdot|$ is the Euclidean norm on $\mathbb{C}^n$. Now, AB has columns $Ab_1, \ldots, Ab_n$, and so

$$\|AB\|_{HS}^2 = \sum_{j=1}^n |Ab_j|^2 \leq \|A\|_{op}^2 \|B\|_{HS}^2,$$

where $\|A\|_{op} = \sup_{|x|=1} |Ax|$ is the operator norm of A; i.e., the largest singular value of A. Writing the singular value decomposition $A = U\Sigma V$ and using the unitary invariance of the Hilbert–Schmidt norm,

$$\|A\|_{op}^2 = \sigma_1^2 \leq \sum_{j=1}^n \sigma_j^2 = \|\Sigma\|_{HS}^2 = \|A\|_{HS}^2,$$

from which the submultiplicativity follows. Indeed, the sharper estimate

$$\|AB\|_{HS} \leq \|A\|_{op}\|B\|_{HS}$$

is often useful.

The discussion above gives two notions of distance on the classical compact matrix groups: first, the Hilbert–Schmidt inner product can be used to define the distance between two matrices A and B by

$$d_{HS}(A,B) := \|A-B\|_{HS} := \sqrt{\langle A-B, A-B\rangle_{HS}} = \sqrt{\mathrm{Tr}\left[(A-B)(A-B)^*\right]}. \tag{1.5}$$

Alternatively, since, for example, $A, B \in \mathbb{U}(n)$ can be thought of as living in a submanifold of Euclidean space $M_n(\mathbb{C})$, one can consider the *geodesic distance* $d_g(A,B)$ between A and B; that is, the length, as measured by the Hilbert–Schmidt metric, of the shortest path lying entirely in $\mathbb{U}(n)$ between A and B. In the case of $\mathbb{U}(1)$, this is arc-length distance, whereas the Hilbert–Schmidt distance defined in Equation (1.5) is the straight-line distance between two points on the circle. Ultimately, the choice of metric is not terribly important:

Lemma 1.3 *Let $A, B \in \mathbb{U}(n)$. Then*

$$d_{HS}(A,B) \leq d_g(A,B) \leq \frac{\pi}{2} d_{HS}(A,B).$$

That is, the two notions of distance are equivalent *in a dimension-free way*.

Proof The inequality $d_{HS}(A,B) \leq d_g(A,B)$ follows trivially from the fact that the Hilbert–Schmidt distance is the geodesic distance in Euclidean space.

For the other inequality, first note that $d_g(A,B) \leq \frac{\pi}{2} d_{HS}(A,B)$ for $A, B \in \mathbb{U}(1)$; that is, that arc-length on the circle is bounded above by $\frac{\pi}{2}$ times Euclidean distance.

Next, observe that both $d_{HS}(\cdot,\cdot)$ and $d_g(\cdot,\cdot)$ are translation-invariant; that is, if $U \in \mathbb{U}(n)$, then

$$d_{HS}(UA, UB) = d_{HS}(A,B) \qquad \text{and} \qquad d_g(UA, UB) = d_g(A,B).$$

In the case of the Hilbert–Schmidt distance, this is immediate from the fact that the Hilbert–Schmidt norm is unitarily invariant. For the geodesic distance, translation invariance follows from the fact that, since any matrix $U \in \mathbb{U}(n)$ acts as an isometry of Euclidean space, every path between A and B lying in $\mathbb{U}(n)$ corresponds to a path between UA and UB of the same length, also lying in $\mathbb{U}(n)$.

Now fix $A, B \in \mathbb{U}(n)$ and let $A^{-1}B = U\Lambda U^*$ be the spectral decomposition of $A^{-1}B$. Then for either distance,

$$d(A,B) = d(I_n, A^{-1}B) = d(I_n, U\Lambda U^*) = d(U^*U, \Lambda) = d(I_n, \Lambda),$$

and so it suffices to assume that $A = I_n$ and B is diagonal.

Write $B = \mathbf{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$. Then the length of the path in $\mathbb{U}(n)$ from A to B given by $U(t) := \mathbf{diag}(e^{it\theta_1}, \dots, e^{it\theta_n})$, for $0 \le t \le 1$ is

$$\begin{aligned}\int_0^1 \|U'(t)\|_{HS}\, dt &= \int_0^1 \left\|\mathbf{diag}(i\theta_1 e^{it\theta_1}, \dots, i\theta_n e^{it\theta_n})\right\|_{HS} dt \\ &= \int_0^1 \sqrt{\theta_1^2 + \cdots + \theta_n^2}\, dt \\ &\le \frac{\pi}{2}\int_0^1 \sqrt{|1 - e^{i\theta_1}|^2 + \cdots + |1 - e^{i\theta_n}|^2}\, dt \\ &= \frac{\pi}{2}\left\|I_n - \mathbf{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})\right\|_{HS},\end{aligned}$$

using the fact that

$$\theta^2 = d_g(1, e^{i\theta})^2 \le \frac{\pi^2}{4} d_{HS}(1, e^{i\theta}),$$

as noted above. □

1.2 Haar Measure

The main goal of this book is to answer the broad general question: What is a random orthogonal, unitary, or symplectic matrix like? To do this, a natural probability measure on each of these groups is needed.

Just as the most natural probability measure (i.e., uniform measure) on the circle is defined by rotation invariance, if G is one of the matrix groups defined in the last section, a "uniform random element" of G should be a random $U \in G$ whose distribution is *translation-invariant*; that is, if $M \in G$ is any fixed matrix, then the equality in distribution

$$MU \stackrel{d}{=} UM \stackrel{d}{=} U$$

should be satisfied. Phrased slightly differently, the distribution of a uniform random element of G should be a translation-invariant probability measure μ on G: for any measurable subset $\mathcal{A} \subseteq G$ and any fixed $M \in G$,

$$\mu(M\mathcal{A}) = \mu(\mathcal{A}M) = \mu(\mathcal{A}),$$

where $M\mathcal{A} := \{MU : U \in \mathcal{A}\}$ and $\mathcal{A}M := \{UM : U \in \mathcal{A}\}$.

It is a theorem due to A. Haar that there is one, and only one, way to do this.

Theorem 1.4 *Let G be any of $\mathbb{O}(n)$, $\mathbb{SO}(n)$, $\mathbb{U}(n)$, $\mathbb{SU}(n)$, or $\mathbb{Sp}(2n)$. Then there is a unique translation-invariant probability measure (called **Haar measure**) on G.*

The theorem is true in much more generality (in particular, any compact Lie group has a Haar probability measure). In the most general case the property of left-invariance is not equivalent to that of right-invariance, but in the case of compact Lie groups, left-invariance implies right-invariance and vice versa, so the phrase "translation invariance" will be used in what follows, and will be assumed to include both left- and right-invariance.

Exercise 1.5

1. Prove that a translation-invariant probability measure on $\mathbb{O}(n)$ is invariant under transposition: if U is Haar-distributed, so is U^T.
2. Prove that a translation-invariant probability measure on $\mathbb{U}(n)$ is invariant under transposition and under conjugation: if U is Haar-distributed, so are both U^T and U^*.

Theorem 1.4 is an existence theorem that does not itself provide a description of Haar measure in specific cases. In the case of the circle, i.e., $\mathbb{U}(1)$, it is clear that Haar measure is just (normalized) arc-length. The remainder of this section gives six different constructions of Haar measure on $\mathbb{O}(n)$, with some comments about adapting the constructions to the other groups. For most of the constructions, the resulting measure is only shown to be invariant one one side; the invariance on the other side then follows from the general fact mentioned above that on compact Lie groups, one-sided invariance implies invariance on both sides.

The Riemannian Perspective

It has already been noted that $\mathbb{O}(n) \subseteq \mathrm{M}_n(\mathbb{R})$ and that it is a compact submanifold. It has two connected components: $\mathbb{SO}(n)$ and $\mathbb{SO}^-(n)$, the set of orthogonal matrices U with $\det(U) = -1$. At each point U of $\mathbb{O}(n)$, there is a tangent space $T_U(\mathbb{O}(n))$, consisting of all the tangent vectors to $\mathbb{O}(n)$ based at U.

A map between manifolds induces a map between tangent spaces as follows. Let M_1, M_2 be manifolds and $\varphi : M_1 \to M_2$. If $x \in T_pM_1$, then there is a curve $\gamma : [0,1] \to M_1$ such that $\gamma(0) = p$ and $\gamma'(0) = x$. Then $\varphi \circ \gamma$ is a curve in M_2 with $\varphi \circ \gamma(0) = \varphi(p)$, and $(\varphi \circ \gamma)'(0)$ is a tangent vector to M_2 at $\varphi(p)$. We take this to be the definition of $\varphi_*(x)$ (it must of course be checked that this gives a well-defined linear map on T_pM_1 for each p).

A Riemannian metric g on a manifold M is a family of inner products, one on the tangent space T_pM to M at each point $p \in M$. The submanifold $\mathbb{O}(n)$ inherits such a metric from $\mathrm{M}_n(\mathbb{R})$, since at each point U in $\mathbb{O}(n)$, $T_U(\mathbb{O}(n))$ is a subspace of $T_U(\mathrm{M}_n(\mathbb{R})) \cong \mathrm{M}_n(\mathbb{R})$. Because multiplication by a fixed

orthogonal matrix V is an isometry of $\mathrm{M}_n(\mathbb{R})$, the induced map on tangent spaces is also an isometry: if $U \in \mathbb{O}(n)$ with $X_1, X_2 \in T_U(\mathbb{O}(n))$ tangent vectors to $\mathbb{O}(n)$ at U, and $R_V : \mathbb{O}(n) \to \mathbb{O}(n)$ denotes multiplication by a fixed $V \in \mathbb{O}(n)$, then

$$g_{VU}((R_V)_* X_1, (R_V)_* X_2) = g_U(X_1, X_2).$$

On any Riemannian manifold, the Riemannian metric uniquely defines a notion of volume. Since the metric is translation-invariant, the normalized volume form on $\mathbb{O}(n)$ is a translation-invariant probability measure; that is, it is Haar measure.

Since each of the classical compact matrix groups is canonically embedded in Euclidean space, this construction works the same way in all cases.

An Explicit Geometric Construction

Recall that $U \in \mathbb{O}(n)$ if and only if its columns are orthonormal. One way to construct Haar measure on $\mathbb{O}(n)$ is to add entries to an empty matrix column by column (or row by row), as follows. First choose a random vector u_1 uniformly from the sphere $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ (that is, according to the probability measure defined by normalized surface area). Take u_1 as the first column of the matrix; by construction, $\|u_1\| = 1$. Now choose u_2 randomly according to surface area measure on

$$\left(u_1^{\perp}\right) \cap \mathbb{S}^{n-1} = \left\{x \in \mathbb{R}^n : \|x\| = 1, \langle x, u_1 \rangle = 0\right\}$$

and let this be the second column of the matrix. Continue in this way; each column is chosen uniformly from the unit sphere of vectors that are orthogonal to each of the preceding columns. The resulting matrix

$$\begin{bmatrix} | & & | \\ u_1 & \dots & u_n \\ | & & | \end{bmatrix}$$

is obviously orthogonal; the proof that its distribution is translation-invariant is as follows.

Observe that if M is a fixed orthogonal matrix, then since

$$M \begin{bmatrix} | & & | \\ u_1 & \dots & u_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ Mu_1 & \dots & Mu_n \\ | & & | \end{bmatrix},$$

the first column of $M \begin{bmatrix} | & & | \\ u_1 & \dots & u_n \\ | & & | \end{bmatrix}$ is constructed by choosing u_1 uniformly from $\mathbb{S}^{n-1}$ and then multiplying by M. But $M \in \mathbb{O}(n)$ means that M acts as a linear isometry of $\mathbb{R}^n$, so it preserves surface area measure on $\mathbb{S}^{n-1}$. That is, the distribution of Mu_1 is exactly uniform on $\mathbb{S}^{n-1}$.

Now, since M is an isometry, $\langle Mu_2, Mu_1\rangle = 0$, and because M is an isometry of $\mathbb{R}^n$, it follows that Mu_2 is *uniformly distributed* on

$$(Mu_1)^\perp \cap \mathbb{S}^{n-1} := \left\{x \in \mathbb{R}^n : |x| = 1, \langle Mu_1, x\rangle = 0\right\}.$$

So the second column of $M\begin{bmatrix} u_1 \dots u_n \end{bmatrix}$ is distributed uniformly in the unit sphere of the orthogonal complement of the first column.

Continuing the argument, the distribution of $M\begin{bmatrix} u_1 \dots u_n \end{bmatrix}$ is exactly the same as the distribution of $\begin{bmatrix} u_1 \dots u_n \end{bmatrix}$; i.e., the construction is left-invariant. It follows by uniqueness that it produces Haar measure on $\mathbb{O}(n)$.

To construct Haar measure on $\mathbb{U}(n)$, one need only draw the columns uniformly from complex spheres in $\mathbb{C}^n$. To get a random matrix in $\mathbb{SO}(n)$, the construction is identical except that there is no choice about the last column; the same is true for $\mathbb{SU}(n)$.

The analogous construction on the representation of elements of $\mathbb{Sp}(2n)$ by $2n \times 2n$ unitary matrices works as follows. For U to be in $\mathbb{Sp}(2n)$, its first column u_1 must lie in the set

$$\{x \in \mathbb{C}^{2n} : \|x\| = 1, \langle x, Jx\rangle = 0\},$$

where J is the matrix defined in (1.4). This condition $\langle x, Jx\rangle = 0$ defines a hyperboloid in $\mathbb{C}^n$ (J is unitarily diagonalizable and has eigenvalues i and $-i$, each with multiplicity n). The set above is thus the intersection of the sphere with this hyperboloid; it is an $(n-2)$-dimensional submanifold of $\mathbb{C}^n$ from which we can choose a point uniformly: this is how we choose u_1. If $n > 1$, one then chooses the second column uniformly from the set

$$\{x \in \mathbb{C}^{2n} : \|x\| = 1, \langle x, u_1\rangle = 0, \langle x, Jx\rangle = 0, \langle x, Ju_1\rangle = 0\};$$

for $n = 1$, one chooses the second column uniformly from

$$\{x \in \mathbb{C}^2 : \|x\| = 1, \langle x, u_1\rangle = 0, \langle x, Jx\rangle = 0\, \langle x, Ju_1\rangle = -1\}.$$

The construction continues: the kth column u_k is chosen uniformly from the intersection of the unit sphere, the hyperboloid $\{x : \langle x, Jx\rangle = 0\}$, and the (affine) subspaces given by the conditions $\langle x, Ju_\ell\rangle = 0$ for $1 \le \ell \le \min\{k-1, n\}$ and $\langle x, Ju_\ell\rangle = -1$ for $n+1 \le \ell < k$ (if $k \ge n+2$). The argument that this