# Entries of Random Matrices

Benjamin Kraft
Liberty High School
Bethlehem, PA, USA

under the direction of
Mr. Gregory Minton
Massachusetts Institute of Technology
Cambridge, MA, USA

Research Science Institute
August 30, 2010

**Entries of Random Matrices**

**Abstract**

Let $U_n$ be the group of $n \times n$ unitary matrices. To select a random unitary matrix, we use the Haar measure. Much study has been devoted to the eigenvalues of random unitary matrices, but little is known about the entries of random unitary matrices and their powers. In this work, we use eigenvalues to understand the entries of random unitary matrices and their powers. We characterize the exact distribution of the top-left entry in the case where the matrix is raised to a power at least $n$, and give some relationships for smaller powers. These results may have applications in quantum mechanics, telephone encryption, and statistical analysis, in addition to helping illuminate the field of random matrix theory.

# 1 Introduction

An orthogonal matrix $A$ is a matrix with real-valued entries such that $AA^{\mathrm{T}} = I$. Similarly, unitary matrices are those with complex-valued entries such that $AA^* = I$, where $A^*$ denotes the conjugate transpose of $A$, that is, if $A$ is $n \times n$, then $A^*_{ij} = \overline{A_{ji}}$ for all $1 \leq i, j \leq n$. In each case, the rows and columns of the matrix are orthonormal vectors, or vectors whose lengths are 1 and whose pairwise dot products are 0. Let $O_n$ be the set of $n \times n$ orthogonal matrices, and $U_n$ be the set of $n \times n$ unitary matrices. Both are compact topological groups. To define a *random* orthogonal or unitary matrix, we use the concept of the Haar measure: in any compact group $G$, it is the unique probability measure $P$ which is translation invariant, that is, for any measurable set $A \subset G$ and any element $M \in G$, $P(A) = P(MA)$.

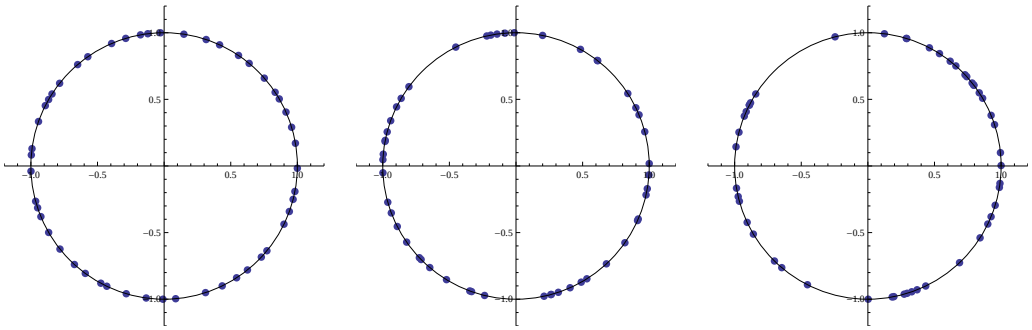
To construct a random Haar distributed orthogonal matrix, it suffices to independently choose each entry of the matrix from a normal distribution centered at zero, then orthonormalize the matrix by the Gram-Schmidt algorithm; see Diaconis [3]. An analogous method also works in the unitary case, using normally distributed complex entries.

For any given matrix $A$, if $A\vec{x} = \lambda\vec{x}$ for some nonzero vector $\vec{x}$ and scalar $\lambda$, then $\lambda$ is called an *eigenvalue* of $A$, and $\vec{x}$ is called an *eigenvector*. The eigenvalues of unitary and orthogonal matrices lie on the unit circle in the complex plane, with those of orthogonal matrices in conjugate pairs. Much study has been devoted to the eigenvalues of random unitary and orthogonal matrices. Diaconis [3] notes several somewhat surprising facts about eigenvalues. First, the eigenvalues tend to repel each other; the probability density for unitary eigenvalues $e^{i\theta_j}$

$$f_2(\theta_1, \ldots, \theta_n) = \frac{1}{(2\pi)^n n!} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2$$

is smaller when two eigenvalues are close together on the unit circle. A derivation is given by Goodman and Wallach [6]. This means the eigenvalues are approximately evenly distributed but otherwise random, as shown in Figure 1. As the matrix is raised to higher powers, the

effect decreases and the eigenvalues become more independent, so that the values tend to clump more, as they have an equal probability of landing near another eigenvalue as far from the other eigenvalues. When a random $n \times n$ unitary matrix $A$ is raised to a power $p \geq n$, the eigenvalues of the resulting matrix are independently and uniformly distributed on the unit circle.



(a) (b) (c)

Figure 1: For comparison we show (a) eigenvalues of $A \in U_{50}$, (b) eigenvalues of $A^{10}$, and (c) 50 independent and uniform $e^{i\theta_j}$. Note that the independent values tend to clump to some extent, while the eigenvalues are approximately evenly spaced. The eigenvalues of $A^{10}$ are in between the two – they clump more than in (a) but less than in (c). The eigenvalues of $A^{50}$ are completely independent and thus look like (c).

Little study has been devoted to the *entries* of random matrices. One of the few known results is due to Borel [1]: if $\Gamma$ is a random orthogonal matrix in $O_n$, then as $n$ gets large, the density of $\sqrt{n}\Gamma_{11}$ approaches $e^{-t^2/2}$. However, not much is known about the exact distributions of the powers of random unitary matrices. We study these entries.

The field of random matrices has many applications in mathematics, along with its own intrinsic value. It is clear that random matrices have a very deep and rich structure which is not well understood. While many specific results are known about random matrices, little is known about the underlying reasons that they behave as they do. This suggests that studying random matrices may be interesting for its own sake. Within mathematics, the zeroes of the Riemann Zeta Function repel each other similarly to the eigenvalues of random unitary matrices. An empirical study by Coram and Diaconis [2] showed that the distribution

of the roots of the Zeta Function was just like that of the eigenvalues.

In addition to connections within mathematics, random matrices have applications to real-world problems. Telephone encryption uses random matrices to encrypt data, and a better understanding of random matrices makes it possible to easily generate pseudorandom matrices which behave like truly random ones. Efficient algorithms using random matrix theory were discovered by Diaconis and Shahshahani [4], Rosenthal [13], and Porod [10]. In statistics, the analysis of large data sets relies on understanding the eigenvalues of large orthogonal matrices; see Mardia, Kent, and Bibby [7]. Finally, random matrices have applications in quantum mechanics, where raising a matrix to a power corresponds to applying the same action repeatedly to a particular quantum state. A sample of such applications is given in Timberlake [16].

In Section 2 we summarize our main results and provide histograms of realizations of these theorems. In Section 3 we establish tools which will be useful in Sections 4 and 5. In Section 4 we consider the case where the matrix is raised to a power at least equal to its dimension, and in Section 5 we consider lower powers.

## 2 Summary of main results

To facilitate the visualization of the results, we include in Figure 2 histograms generated from simulations of the first entries of 20,000 random matrices from $U_n$ raised to the $p^{\text{th}}$ power for $n = 3$, 4, and 5 and $p = 1, 2, \ldots, 5$. For the $p = 3$ and $n = 3$ case a three-dimensional histogram of 1,000,000 random matrices is shown in Figure 3.

Our first theorem deals with the distribution of entries of high powers of random unitary matrices. Let $(A^p)_{11}$ be the top left element of $A^p$. Then when $p \geq n$, if we select $A$ randomly from $U_n$, the distribution of $(A^p)_{11}$ becomes the same for all $p \geq n$. We give a density function for that distribution.

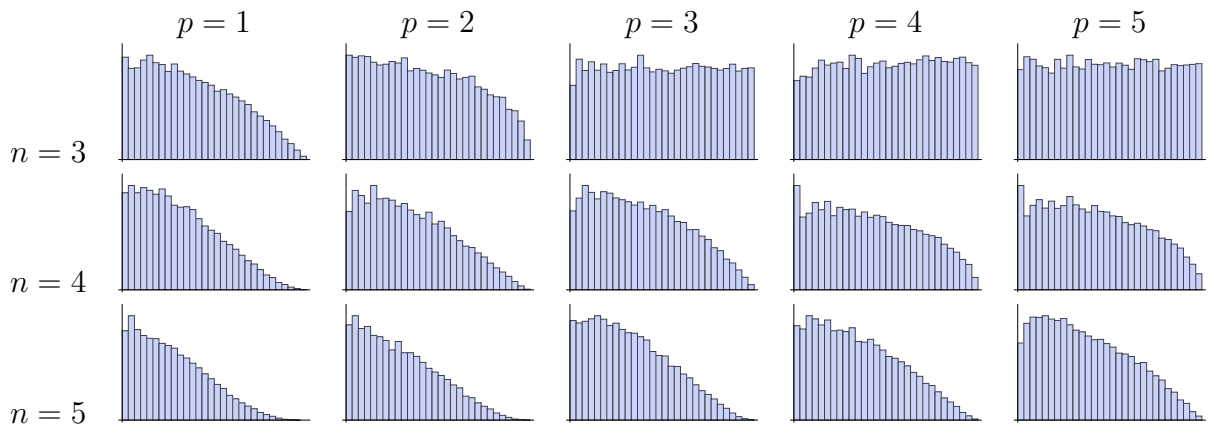



Figure 2: Histograms of a cross-section of the distribution of $(U^p)_{11}$ on the unit circle, generated by considering the magnitudes of $(U^p)_{11}$, then normalizing for the increased density of points towards the edge of the unit disk. The horizontal axis goes from 0 to 1, that is, from the center of the disk to the edge, and the vertical axis is the density. Note, for example, that when $p \geq n$, the probability density stabilizes, and that for $n = 3$ and $p \geq 3$, the density is uniform (since by Theorem 1 the density function is constant).

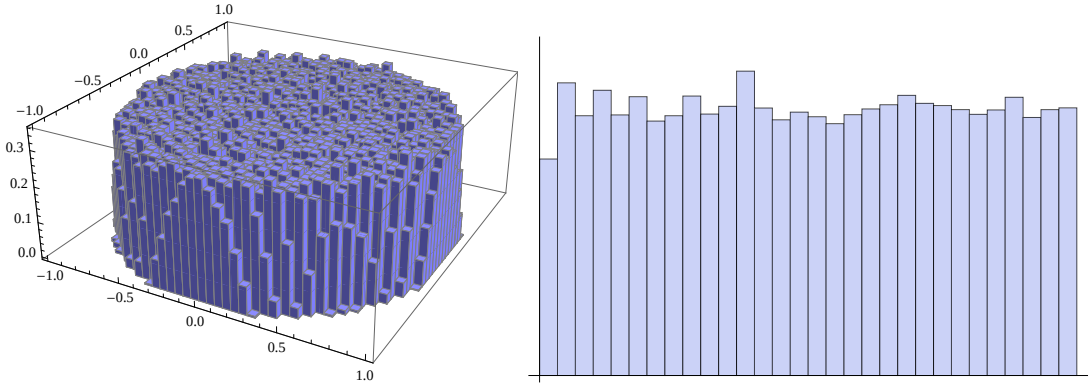



Figure 3: A three-dimensional histogram of a simulation with 1,000,000 matrices with $n = 3$ and $p = 3$. The $x$ and $y$ axes are the real and imaginary parts of the first entry, and the $z$ axis is the probability density. Notice that the distribution is uniform and isotropic, just as shown in the cross-section histograms and just as predicted by Theorem 1. For comparison, the corresponding cross-section histogram is included.

**Theorem 1.** *Let $n \geq 2$, and let $A$ be a random $n \times n$ unitary matrix selected according to Haar measure. Then $Z_n = (A^p)_{11}$ has probability density function $f_{Z_n}(\zeta) = c(1 - |\zeta|^2)^{\frac{n-3}{2}}$ on the unit disk for all integers $p \geq n$, where $c = \frac{n-1}{2\pi}$.*

Our second theorem deals with the distribution of lower powers of random unitary matrices. When $1 < p < n$, the exact distribution of the first entry does not appear to have a simple density function. However, when $\lceil \frac{n}{2} \rceil \leq p \leq n$, some information can be gained about the moments of the squared magnitudes of the distributions.

**Theorem 2.** *Let $n \geq 2$, and let $\lceil \frac{n}{2} \rceil \leq p_1 < p_2 \leq n$. Let $A$ be a random $n \times n$ unitary matrix selected according to Haar measure. If the $e^{th}$ moment of $X = |(A^{p_1})_{11}|^2$ is $\mu_e$, and the $e^{th}$ moment of $Y = |(A^{p_2})_{11}|^2$ is $\nu_e$, then when $e$ is a positive integer, $\mu_e < \nu_e$.*

In effect, this means that as $p$ increases, $Z = (A^p)_{11}$ is more likely to fall towards the edge of the unit disk than it is for $p$ small. For example, when $n = 3$, $(A^2)_{11}$ is distributed mostly towards the center of the disk, while $(A^3)_{11}$ is uniform over the disk by Theorem 1. This can also be observed in Figure 2, where the probability densities bulge more towards the right as $p$ gets larger.

We conjecture that Theorem 2 holds for all $1 \leq p_1 < p_2 \leq n$; statistical samples appear to support this conjecture.

# 3 Tools for understanding random matrices

One important tool in the study of random matrices is that of moments. Given any continuous probability distribution $X$ over $\mathbb{R}$ with density function $f$, the $e^{\text{th}}$ moment of $X$ is defined as

$$\mu_e = E[X^e] = \int_{-\infty}^{\infty} x^e f(x)\, dx.$$

**Remark 3** (Feller [5])**.** *The moments of a bounded distribution uniquely define that distribution.*

*Proof.* In brief, the moments determine the distribution since they determine the Fourier Transform of the density function, which in turn determines the distribution. A full proof appears on page 233 in Feller [5]. □

While finding the density function from the moments may be difficult, understanding the moments can be key to understanding the probability distribution.

Let $\Re(z)$ be the real part of the complex number $z$. We note a consequence of a classical theorem of probability theory.

**Remark 4.** *If $Z$ is a distribution over $\mathbb{C}$ which is independent of phase, then $\Re(Z)$ uniquely determines $Z$.*

*Proof.* It is a classical theorem of probability theory that a probability measure on $\mathbb{R}^n$ is completely determined by its 1-dimensional marginals (a proof is given as Remark 2.3.5 in Stroock [15]). The result is merely a special case of this fact. □

Next, we derive Lemma 5, a well-known integral identity which is useful in the proof of Theorem 1.

**Lemma 5.** *For all even nonnegative integers $k$,*

$$\int_0^{2\pi} \cos(\phi)^k \, d\phi = 2\sqrt{\pi}\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k+2}{2}\right)}.$$

*Proof.* We induct on $k$. If $k = 0$, then

$$\int_0^{2\pi} \cos(\phi)^0 \, d\phi = \int_0^{2\pi} d\phi = 2\pi = 2\sqrt{\pi}\frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma(1)}.$$

Now suppose that

$$\int_0^{2\pi} \cos(\phi)^{k-2} \, d\phi = 2\sqrt{\pi}\frac{\Gamma\left(\frac{k-1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}.$$

Then integrating by parts with $u = \cos(\phi)^{k-1}$ and $dv = \cos(\phi)\,d\phi$ gives

$$\begin{aligned}
\int_0^{2\pi} \cos(\phi)^k\,d\phi &= \cos(\phi)^{k-1}\sin(\phi)\Big|_0^{2\pi} - \int_0^{2\pi} (k-1)\cos(\phi)^{k-2}(-\sin(\phi))\sin(\phi)\,d\phi \\
\int_0^{2\pi} \cos(\phi)^k\,d\phi &= (k-1)\int_0^{2\pi} \cos(\phi)^{k-2}\,d\phi - (k-1)\int_0^{2\pi} \cos(\phi)^k\,d\phi \\
k\int_0^{2\pi} \cos(\phi)^k\,d\phi &= (k-1)\int_0^{2\pi} \cos(\phi)^{k-2}\,d\phi \\
\int_0^{2\pi} \cos(\phi)^k\,d\phi &= \frac{k-1}{k} 2\sqrt{\pi}\frac{\Gamma\left(\frac{k-1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} = 2\sqrt{\pi}\frac{\frac{k-1}{2}\Gamma\left(\frac{k-1}{2}\right)}{\frac{k}{2}\Gamma\left(\frac{k}{2}\right)} = 2\sqrt{\pi}\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k+2}{2}\right)}.
\end{aligned}$$

□

Finally, Lemma 6, another well-known lemma about trigonometric integrals, is useful in the proof of Theorem 2.

**Lemma 6.** *If $a$ and $b$ are nonnegative integers, then $\int_0^{2\pi}(\sin\theta)^a(\cos\theta)^b\,d\theta$ is zero if at least one of $a$ or $b$ is odd, and positive otherwise.*

*Proof.* We merely use the symmetry of the sine and cosine functions.

$$\begin{aligned}
\int_0^{2\pi}(\sin\theta)^a(\cos\theta)^b\,d\theta &= \int_0^{\pi/2}(\sin\theta)^a(\cos\theta)^b\,d\theta + \int_{\pi/2}^{\pi}(\sin\theta)^a(\cos\theta)^b\,d\theta \\
&\quad + \int_{\pi}^{3\pi/2}(\sin\theta)^a(\cos\theta)^b\,d\theta + \int_{3\pi/2}^{\pi}(\sin\theta)^a(\cos\theta)^b\,d\theta \\
\int_0^{2\pi}(\sin\theta)^a(\cos\theta)^b\,d\theta &= \int_0^{\pi/2}(\sin\theta)^a(\cos\theta)^b\,d\theta + (-1)^b\int_0^{\pi/2}(\sin\theta)^a(\cos\theta)^b\,d\theta \\
&\quad + (-1)^{a+b}\int_0^{\pi/2}(\sin\theta)^a(\cos\theta)^b\,d\theta + (-1)^b\int_0^{\pi/2}(\sin\theta)^a(\cos\theta)^b\,d\theta.
\end{aligned} \tag{1}$$

When at least one of $a$ or $b$ is odd, exactly two of the powers of $-1$ in (1) are odd powers, so the whole integral is zero. When both are even, (1) simplifies to

$$4\int_0^{\pi/2}(\sin\theta)^a(\cos\theta)^b\,d\theta,$$

which is always greater than zero. □

# 4 Higher Powers

We now proceed to the proofs of our main theorems. In the ensuing discussion of powers of random unitary matrices, it is important to understand random (real) unit vectors on the $n$-dimensional sphere. One way to generate such a random vector is to generate $n$ random and normally distributed real numbers, then divide each by the sum of their squares. However, it is often much simpler to use the distribution of a single entry. Therefore, we begin with a lemma about this distribution. Throughout the paper, we let $c$ be a normalization constant independent of the moment number.

**Lemma 7.** *If $\vec{X} = (X_1, X_2, \ldots, X_n)$ is a random unit vector in $n$ dimensions, then the density function for $X_1$ is $f_{X_1}(x) = c(1-x^2)^{\frac{n-3}{2}}$ on $[-1, 1]$.*

The proof is a common exercise in multivariable calculus and is given in Appendix A.

**Lemma 8.** *Let $X$ be the first coordinate of a random real unit vector in $2n$ dimensions, and let $Y$ be the first coordinate of an independent random real unit vector in $2n-1$ dimensions. Then $W = (1-X^2)(1-Y^2)$ has density function $f_W(w) = cw^{n-2}$ on $[0, 1]$.*

*Proof.* By Lemma 7, $X$ has density function $f_X(x) = c(1-x^2)^{\frac{2n-3}{2}}$ and $Y$ has density function $f_Y(y) = c(1-y^2)^{\frac{2n-4}{2}}$. Then if $B(x, y)$ is the beta function, the $e^{\text{th}}$ moment of $W$ is

$$\begin{aligned}
\mu_e &= E[(1-X^2)^e(1-Y^2)^e] \\
&= c\int_{-1}^{1}\int_{-1}^{1}(1-x^2)^e(1-y^2)^e(1-x^2)^{\frac{2n-3}{2}}(1-y^2)^{\frac{2n-4}{2}}\,dy\,dx \\
&= c\left(\frac{1}{2}\int_{-1}^{1}x^{-1}(1-x^2)^{e+n-\frac{3}{2}}2x\,dx\right)\left(\frac{1}{2}\int_{-1}^{1}y^{-1}(1-y^2)^{e+n-2}2y\,dy\right) \\
&= c\left(\int_0^1 u^{-\frac{1}{2}}(1-u)^{e+n-\frac{3}{2}}\,du\right)\left(\int_0^1 v^{-\frac{1}{2}}(1-v)^{e+n-2}\,dv\right) \qquad (u=x^2, v=y^2) \\
&= cB\left(\frac{1}{2}, e+n-\frac{1}{2}\right)B\left(\frac{1}{2}, e+n-1\right) \\
&= c\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(e+n-\frac{1}{2}\right)}{\Gamma(e+n)}\frac{\Gamma\left(\frac{1}{2}\right)\Gamma(e+n-1)}{\Gamma\left(e+n-\frac{1}{2}\right)}
\end{aligned}$$

$$\mu_e = c\frac{1}{e+n-1},$$

where the $\Gamma\left(\frac{1}{2}\right)^2$ is absorbed into $c$. But $\mu_0 = E[Z^0] = E[1] = 1$ so $c = n-1$. Then $\mu_e = \frac{n-1}{e+n-1}$.

Now consider the moments $\nu_e$ of the random variable $V$ with density function $f_V(v) = cv^{n-2}$ on $[0,1]$:

$$\nu_e = E[V^e] = c\int_0^1 v^e v^{n-2}\,dx = c\frac{1}{e+n-1}.$$

Again, $\nu_0 = 1$ so $c = n-1$. Then $\nu_e = \frac{n-1}{e+n-1}$. The moments of $V$ are identical to those of $W$, so since both distributions are bounded (in $[0,1]$), they must be the same, that is, $f_W(w) = cw^{n-2}$. □

Remark 4 suggests that understanding the relationships between the density function of a distribution and the density function of its real part is important. The next lemma gives a specific relation.

**Lemma 9.** *If $Z$ is a random variable with density function $f_Z(\zeta) = c(1-|\zeta|^2)^k$ on the unit disk, then the density of its real part is $f_{\Re(Z)}(x) = c(1-x^2)^{k+\frac{1}{2}}$ on $[-1,1]$.*

*Proof.* We consider the moments of $\Re(Z)$. Let $Z = X + iY$ and consider the $e^{\text{th}}$ moment of $\Re(Z)$.

$$\begin{aligned}
E[\Re(Z)^e] &= c\int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{-\sqrt{1-x^2}} \Re(x+iy)^e(1-|x+iy|^2)^k\,dy\,dx \\
&= c\int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} x^e(1-x^2-y^2)^k\,dy\,dx \\
&= c\int_{-1}^{1} x^e(1-x^2)^{k+\frac{1}{2}}\left(2\int_0^{\sqrt{1-x^2}}\frac{\sqrt{1-x^2}}{y}\left(1-\frac{y^2}{1-x^2}\right)^k\frac{y}{1-x^2}\,dy\right)dx \\
&= c\int_{-1}^{1} x^e(1-x^2)^{k+\frac{1}{2}}\left(\int_0^1 u^{-\frac{1}{2}}(1-u)^k\,du\right)dx \qquad \left(\text{where } u = \frac{y^2}{1-x^2}\right) \\
&= c\int_{-1}^{1} x^e(1-x^2)^{k+\frac{1}{2}}B\left(\frac{1}{2},k+1\right)dx \\
&= c\int_{-1}^{1} x^e(1-x^2)^{k+\frac{1}{2}}\,dx,
\end{aligned}$$

where the beta function is independent of the moment number and thus becomes part of $c$. But these are just the moments of the distribution with density function $f(x) = c(1-x^2)^{k+\frac{1}{2}}$, so the lemma is complete. □

We now proceed to the proof of Theorem 1. Let $U$ be a random $n \times n$ unitary matrix selected according to Haar measure, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be its eigenvalues corresponding to eigenvectors $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}$. Then since unitary matrices are diagonalizable, if $Q$ is the matrix whose columns are the $\vec{v_i}$, and $D$ is the diagonal matrix with the $\lambda_i$ along the diagonal, we know that $U = Q^{-1}DQ$, from which we can see that $U^p = Q^{-1}D^pQ$. In addition, Marzetta et al. [8] proved that the eigenvectors of $U$ are isotropic and independent of the eigenvalues, so $Q$ is a random unitary matrix independent of $D$. Note that the eigenvalues of $U^p$ are exactly the $\lambda_i^p$ which appear along the diagonal of $D^p$. Rains [11] proved that whenever $p \geq n$, the eigenvalues of $U^p$ are distributed as $n$ points chosen independently and uniformly on the unit circle. This allows us to find the entries of the corresponding powers of $U$.

**Theorem 1.** *Let $n \geq 2$, and let $A$ be a random $n \times n$ unitary matrix selected according to Haar measure. Then $Z_n = (A^p)_{11}$ has probability density function $f_{Z_n}(\zeta) = c(1-|\zeta|^2)^{\frac{n-3}{2}}$ on the unit disk for all integers $p \geq n$, where $c = \frac{n-1}{2\pi}$.*

*Proof.* Let $U = Q^{-1}DQ = Q^*DQ$ where $Q^*$ represents the conjugate transpose of $Q$, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of $U$ corresponding to eigenvectors $\vec{v_i}$. Then

$$Z_n = (U^p)_{11} = \begin{pmatrix} \overline{Q_{11}} & \overline{Q_{21}} & \cdots & \overline{Q_{n1}} \end{pmatrix} \begin{pmatrix} \lambda_1^p & 0 & \cdots & 0 \\ 0 & \lambda_2^p & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^p \end{pmatrix} \begin{pmatrix} Q_{11} \\ Q_{21} \\ \vdots \\ Q_{n1} \end{pmatrix}$$

$$Z_n = \sum_{i=1}^{n} \lambda_i^p Q_{i1} \overline{Q_{i1}}$$

$$Z_n = \sum_{i=1}^{n} \lambda_i^p |Q_{i1}|^2. \tag{2}$$

However, the $\lambda_i^p$ are independent and uniform on the unit circle, so we may replace them with random variables $e^{i\Phi_i}$ which are independent and uniform on the unit circle, eliminating the eigenvalues from the equation. In addition, $Q$ is a random unitary matrix independent of the $\lambda_i$ and thus the $\Phi_i$, so its first column is a random complex unit vector in $n$ dimensions.

We now prove that the real part of $Z_n$ has density function $f_{\Re(Z_n)}(\zeta) = (1-\zeta^2)^{\frac{n-2}{2}}$ by induction on $n$. As a base case, let $n = 1$. Since $Z_1$ is just a uniform distribution over the unit circle, Lemma 7 shows that $f_{\Re(Z_n)}(\zeta) = (1-\zeta^2)^{-\frac{1}{2}}$ is the distribution of the first component, that is, the real part of $Z_1$.

Now suppose that the density function for $\Re(Z_{n-1})$ is $f_{\Re(Z_{n-1})}(\zeta) = c(1-\zeta^2)^{\frac{(n-1)-2}{2}}$. Pulling the first term out of (2) gives

$$Z_n = |Q_{11}|^2 e^{i\Phi} + (1-|Q_{11}|^2)Z_{n-1}$$

$$\Re(Z_n) = |Q_{11}|^2 \Re(e^{i\Phi}) + (1-|Q_{11}|^2)\Re(Z_{n-1}),$$

where $Q_{11}$ is selected as the first coordinate of a random complex $n$-dimensional unit vector and $\Phi$ is uniform over the interval $[0, 2\pi]$; we scale $Z_{n-1}$ by $1-|Q_{11}|^2$ to account for the fact that the remaining $Q_{i1}$ are chosen as a random vector with squared magnitude $1-|Q_{11}|^2$. Switching over to real-valued variables, we select $X$ as the first coordinate of a $2n$-dimensional unit vector, then select $Y$ as the first coordinate of a $(2n-1)$-dimensional unit vector. Then, using a scaling factor since the imaginary part of $Q_{11}$ is in effect selected from a vector of length $\sqrt{1-X^2}$,

$$Q_{11} = X + Yi\sqrt{1-X^2}$$

$$|Q_{11}|^2 = X^2 + (1-X^2)Y^2 = 1 - (1-X^2)(1-Y^2).$$

Therefore, $X$ and $Y$ can be replaced by the single random variable $W$, where $W = (1-X^2)(1-Y^2)$ has distribution $f_W(w) = cw^{n-2}$ by Lemma 8.

We proceed by calculating the moments of $\Re(Z_n)$:

$\mu_e = E[(\Re((1-W)e^{i\Phi} + WZ_{n-1}))^e]$

$$
\begin{aligned}
&= c\int_{-1}^{1}\int_{0}^{2\pi}\int_{0}^{1}((1-w)\cos(\phi)+w\zeta)^e w^{n-2}\frac{1}{2\pi}(1-\zeta^2)^{\frac{n-3}{2}}\,dw\,d\phi\,d\zeta \\
&= c\int_{-1}^{1}\int_{0}^{2\pi}\int_{0}^{1}\left(\sum_{k=0}^{e}\binom{e}{k}(1-w)^k\cos(\phi)^k w^{e-k}\zeta^{e-k}\right)w^{n-2}(1-\zeta^2)^{\frac{n-3}{2}}\,dw\,d\phi\,d\zeta \\
&= c\sum_{k=0}^{e}\binom{e}{k}\left(\int_{-1}^{1}\zeta^{e-k}(1-\zeta^2)^{\frac{n-3}{2}}\,d\zeta\right)\left(\int_{0}^{2\pi}\cos(\phi)^k\,d\phi\right)\left(\int_{0}^{1}(1-w)^k w^{e-k+n-2}\,dw\right).
\end{aligned}
$$

If $e$ is odd, either $k$ is odd, in which case the integral involving $\phi$ is zero, or $k$ is even so $e-k$ is odd, in which case that involving $\zeta$ is zero, thus the entire expression is 0. When $e$ is even, we set $u=\zeta^2$ and get

$$
\begin{aligned}
\mu_e &= c\sum_{\substack{k=0\\ \text{even}}}^{e}\binom{e}{k}\left(\int_{0}^{1}\zeta^{e-k-1}(1-\zeta^2)^{\frac{n-3}{2}}2\zeta\,d\zeta\right)\left(\int_{0}^{2\pi}\cos(\phi)^k\,d\phi\right)\left(\int_{0}^{1}(1-w)^k w^{e-k+n-2}\,dw\right) \\
&= c\sum_{\substack{k=0\\ \text{even}}}^{e}\binom{e}{k}\left(\int_{0}^{1}u^{\frac{e-k-1}{2}}(1-u)^{\frac{n-3}{2}}\,du\right)\left(\int_{0}^{2\pi}\cos(\phi)^k\,d\phi\right)\left(\int_{0}^{1}(1-w)^k w^{e-k+n-2}\,dw\right) \\
&= c\sum_{\substack{k=0\\ \text{even}}}^{e}\binom{e}{k}B\left(\frac{e-k+1}{2},\frac{n-1}{2}\right)\left(\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k+2}{2}\right)}\right)B(e-k+n-1,k+1) \\
\mu_e &= c\sum_{\substack{k=0\\ \text{even}}}^{e}\binom{e}{k}\left(\frac{\Gamma\left(\frac{e-k+1}{2}\right)\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{e-k+n}{2}\right)}\right)\left(\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k+2}{2}\right)}\right)\left(\frac{\Gamma\left(e-k+n-1\right)\Gamma\left(k+1\right)}{\Gamma\left(e+n\right)}\right).
\end{aligned}
$$

It can be verified (see Appendix B) that, up to constant factors dependent only on $n$, this is equal to

$$
\mu_e = c\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{e+1}{2}\right)}{\Gamma\left(\frac{n+e+1}{2}\right)}.
$$

Now we compute the moments $\nu_e$ of desired distribution whose real part has density function $f(x)=c(1-x^2)^{\frac{n-2}{2}}$.

$$
\begin{aligned}
\nu_e &= c\int_{-1}^{1}x^e(1-x^2)^{\frac{n-2}{2}}\,dx \\
&= c\int_{0}^{1}u^{\frac{e-1}{2}}(1-u)^{\frac{n-3}{2}}\,du \qquad\qquad (\text{with } u=x^2) \\
&= cB\left(\frac{e+1}{2},\frac{n}{2}\right)
\end{aligned}
$$

$$= c\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{e+1}{2}\right)}{\Gamma\left(\frac{n+e+1}{2}\right)}$$

$$\nu_e = \mu_e,$$

where the constants must be the same since $\nu_0 = 1 = \mu_0$. Then the real parts of the distributions must be the same, so by induction, $f_{\Re(Z_n)}(\zeta) = c(1-\zeta^2)^{\frac{n-2}{2}}$ for all $n$. Thus, by Remark 4 and Lemma 9,

$$f_{Z_n}(\zeta) = c(1-|\zeta|^2)^{\frac{n-3}{2}}$$

for all $n$; the value $c = \frac{n-1}{2\pi}$ is easy to verify. □

# 5 Lower powers

When considering powers less than $n$ of uniformly random matrices, a result of Rains [12] gives an important way of studying the eigenvalues, and through them, the entries of the powers of a uniformly random matrix. Rains showed that the distribution of eigenvalues of the $p^{\text{th}}$ power of a random unitary matrix is equivalent to the direct sum of $p$ independent distributions, each of which is the distribution of a smaller random unitary matrix. Specifically, if $A_n$ is a uniformly random $n \times n$ unitary matrix, then

$$A_n^p \sim \bigoplus_{0 \le i < p} A_{\lceil \frac{n-i}{p} \rceil}.$$

Therefore, understanding the effect of the correlation between eigenvalues, even if only between two, is critical in understanding powers of larger $A_n$. Recall that the probability density for the two eigenvalues of a matrix selected from $U_2$ is

$$\begin{aligned} f(\theta_1, \theta_2) &= c|e^{i\theta_1} - e^{i\theta_2}|^2 \\ &= c((\cos\theta_1 - \cos\theta_2)^2 + (\sin\theta_1 - \sin\theta_2)^2) \\ &= c(2 - 2\cos\theta_1\cos\theta_2 - 2\sin\theta_1\sin\theta_2) \\ &= c(1 - \cos(\theta_1 - \theta_2)), \end{aligned} \tag{3}$$

where the factor of 2 is pulled into the constant. We continue with a lemma involving this

density.

**Lemma 10.** *Let $x_1$ and $x_2$ be fixed positive real numbers, and let $Y_1$ and $Y_2$ be random variables such that $Y_1 = x_1e^{i\theta_1} + x_2e^{i\theta_2}$ where the $\theta_i$ are independent and uniform over $[0, 2\pi]$ and such that $Y_2 = x_1e^{i\theta_1} + x_2e^{i\theta_2}$ where the $\theta_i$ are selected with probability density $c(1-\cos(\theta_1-\theta_2))$, with each $\theta_i$ in $[0, 2\pi]$. Then the positive integral moments of the magnitude of $Y_1$ are greater than the corresponding moments of the magnitude of $Y_2$.*

*Proof.* Let $\mu_e$ be the $e^{\text{th}}$ moment of $Y_1$, and let $\nu_e$ be the $e^{\text{th}}$ moment of $Y_2$. Then

$$\begin{aligned}
\nu_e &= c\int_0^{2\pi}\int_0^{2\pi} |x_1e^{i\theta_1} + x_2e^{i\theta_2}|^e(1-\cos(\theta_1-\theta_2))\,d\theta_1\,d\theta_2 \\
&= c\int_0^{2\pi}\int_{\theta_2}^{2\pi+\theta_2} |x_1e^{i(\theta_1-\theta_2)} + x_2|^e(1-\cos(\theta_1-\theta_2))\,d\theta_1\,d\theta_2 \\
&= c\int_0^{2\pi}\int_0^{2\pi} |x_1e^{i\delta} + x_2|(1-\cos\delta)\,d\delta\,d\theta_2 && \text{(letting } \delta = \theta_1-\theta_2\text{)} \\
&= c\cdot 2\pi\cdot 2\int_0^{\pi} |e^{i\delta} + x|(1-\cos\delta)\,d\delta && \text{(with } x = \frac{x_2}{x_1} > 0\text{)} \\
&= c\left(\int_0^{\pi/2} |e^{i\delta} + x|(1-\cos\delta)\,d\delta + \int_{\pi/2}^{\pi} |e^{i\delta} + x|(1-\cos\delta)\,d\delta\right) \\
&= \mu_e + c\left(\int_0^{\pi/2} |e^{i\delta} + x|(-\cos\delta)\,d\delta + \int_0^{\pi/2} |e^{i(\pi-\delta)} + x|\cos\delta\,d\delta\right) \\
\nu_e &= \mu_e + c\int_0^{\pi/2} (\cos\delta)(|e^{i(\pi-\delta)} + x| - |e^{i\delta} + x|)\,d\delta
\end{aligned}$$

But the real part of $e^{i\delta}$ is greater than that of $e^{i(\pi-\delta)}$ whenever $0 \le \delta < \frac{\pi}{2}$ and the imaginary parts of the two are equal, so the integrand is always negative on $[0, \frac{\pi}{2}]$. Then $\nu_e < \mu_e$ as desired. □

**Theorem 2.** *Let $n \ge 2$, and let $\lceil \frac{n}{2} \rceil \le p_1 < p_2 \le n$. Let $A$ be a random $n \times n$ unitary matrix selected according to Haar measure. Then if the $e^{th}$ moment of $X = |(A^{p_1})_{11}|^2$ is $\mu_e$, and the $e^{th}$ moment of $Y = |(A^{p_2})_{11}|^2$ is $\nu_e$, then when $e$ is a positive integer, $\mu_e < \nu_e$.*

We can now proceed to the proof of Theorem 2.

*Proof.* Let $B_p$ be a random variable representing the top left entry of the $p^{\text{th}}$ power of a random $n \times n$ unitary matrix, and let $\lceil \frac{n}{2} \rceil \leq p < n$. Then we want to show that moments of the squared magnitude of $B_{p+1}$ are greater than those of the squared magnitude of $B_p$. The result follows.

Since $\lceil \frac{n}{2} \rceil \leq p < n$, the eigenvalues of the $p^{\text{th}}$ power of a random $n \times n$ unitary matrix are the union of sets of eigenvalues chosen as those of some number of random $2 \times 2$ matrices along with some number of independent eigenvalues. The eigenvalues of the $(p+1)^{\text{st}}$ power of the same matrix consist of one fewer pair of correlated eigenvalues and two more independent ones, that is, one of the sets of eigenvalues from $U_2$ becomes two independent eigenvalues. Let $B$ be the weighted sum of the unchanged eigenvalues (those selected in the same way in both cases), and let $x_1$ and $x_2$ be the remaining entries in the first eigenvector, such that $B_p = x_1 e^{i\theta_1} + x_2 e^{i\theta_2} + B$, with the $\theta_i$ having density function $f(\theta_1, \theta_2) = c(1 - \cos(\theta_1 - \theta_2))$, and $B_{p+1} = x_1 e^{i\theta_1} + x_2 e^{i\theta_2} + B$ with independent and uniform $\theta_1$ and $\theta_2$. Note that $\theta_1$ and $\theta_2$ are independent of the $x_i$ and $B$. Fix $B$ and the $x_i$, then let $xe^{i\alpha} = \frac{1}{B}(x_1 e^{i\theta_1} + x_2 e^{i\theta_2})$, where $x$ depends on the $\theta_i$ and $\alpha$ is independent and uniform on $[0, 2\pi]$. Note that Lemma 10 means that after $B$ and $x_i$ are fixed, the moments of $x$ are larger when the $\theta_i$ are uncorrelated, in $B_{p+1}$. But then we can calculate the moments of the squared magnitude of $B_{p+1}$:

$$\begin{aligned}
E[|B_{p+1}|^{2e}] &= E[B^e |xe^{i\alpha} + 1|^{2e}] \\
&= B^e E\left[\sum_{k=0}^{e} \binom{e}{k} (x\cos\alpha + 1)^{2k} (\sin\alpha)^{2e-2k}\right] \\
&= B^e E\left[\sum_{k=0}^{e} \binom{e}{k} \sum_{j=0}^{2k} \binom{2k}{j} x^j (\cos\alpha)^j (\sin\alpha)^{2e-2k}\right] \\
&= B^e \sum_{k=0}^{e} \binom{e}{k} \sum_{j=0}^{2k} \binom{2k}{j} E[x^j] E[(\cos\alpha)^j (\sin\alpha)^{2e-2k}].
\end{aligned}$$

By Lemma 6, the second expected value is nonnegative; if $e > 0$, then at least one term must be positive. Then when $E[x^j]$ increases, so does $|B_{p+1}|^{2e}$, so the moments of the squared

magnitude of $B_{p+1}$ are greater than the corresponding moments of the squared magnitude of $B_p$. $\square$

# 6 Conclusion

In the preceding sections, we have explored the distributions of entries of powers of random unitary matrices. We have exactly characterized the distributions of the top left entries of matrices raised to powers at least equal to the dimension. Specifically, if $U$ is an $n \times n$ random unitary matrix, then $(U^p)_{11}$ has density function

$$f(\zeta) = c(1 - |\zeta|)^{\frac{n-3}{2}}$$

over the unit disk. Furthermore, while the distributions of lower powers appear to be difficult to exactly describe, we have given results in specific cases – the moments of the magnitudes of the distributions increase as $p$ increases, whenever $p$ is at least $\lceil \frac{n}{2} \rceil$ and at most $n$. We conjecture that this relationship extends to all $1 \leq p \leq n$.

This work has direct relation to the applications of random matrix theory discussed in the introduction. The quantum mechanical applications in particular are directly related to the first entry of the powers of $U$. However, in a deeper sense, the field of random matrices is not well understood. Any new method of approach gives another possible means to figure out why random matrices really behave the way they do. We hope our work will help illuminate the field as a whole.

# References


[1] E. Borel. Sur les Principes de la Theorie Cinetique des gaz. Annales, L'Ecole Normal Sup. (1906), no. 23, 9–32.

[2] M. Coram and P. Diaconis. New tests of the correspondence between unitary eigenvalues and the zeros of Riemann's zeta function. Journal of Physics A: Mathematical and General 36 (2003), no. 12, 2883–2906.

[3] P. Diaconis. Patterns in eigenvalues. The $70^{\text{th}}$ Josiah Willard Gibbs Lecture. Bulletin of the American Mathematical Society 40 (2003), no. 2, 155–178.

[4] P. Diaconis and M. Shahshahani. Products of random matrices as they arise in the study of random walks on groups. Contemporary Mathematics 50 (1986), 183–195.

[5] W. Feller. *An Introduction to Probability Theory and Its Applications, Vol. 2.* Wiley, Hoboken, NJ (1971).

[6] R. Goodman and W. Wallach. *Representations and Invariants of the Classical Groups.* Cambridge Press, Cambridge (1998).

[7] K. Mardia, J. Kent, and J. Bibby. *Multivariate Analysis.* Academic Press: New York (1979).

[8] T. L. Marzetta, B. Hassibi, and B. M. Hochwald. Structured unitary space-time autocoding constellation. IEEE Transactions on Information Theory 48 (2002), no. 4, 942–950.

[9] P. Paule and M. Schorn. A Mathematica version of Zeilberger's algorithm for proving binomial coefficient identities. Journal of Symbolic Computation 20 (1995), no. 5-6, 673–698.

[10] U. Porod. The cut-off phenomenon for random reflection. Annals of Probability 24 (1996), no. 1, 74–96.

[11] E. Rains. High powers of random elements of compact lie groups. Probabililty Theory and Related Fields 107 (1997), no. 2, 219–241.

[12] E. Rains. Images of Eigenvalue Distributions under Power Maps. Probability Theory and Related Fields 125 (2003), no. 4, 522–538.

[13] J. Rosenthal. Random rotations, characters and random walks on $SO(N)$. Annals of Probability 22 (1994), no. 1, 398-423.

[14] M. Schorn. Contributions to Symbolic Summation. Diploma Thesis, RISC, J. Kepler University, Linz (1995).

[15] D. W. Stroock. *Probability Theory: An Analytic View.* Cambridge University Press: Cambridge, 2000.

[16] T. Timberlake. Random numbers and random matrices: Quantum chaos meets number theory. American Journal of Physics 74 (2006), no. 6, 547–553.

# A Proof of Lemma 7

**Lemma 7.** *If $\vec{X} = (X_1, X_2, \ldots, X_n)$ is a random unit vector in $n$ dimensions, then the density function for $X_1$ is $f_{X_1}(x) = c(1-x^2)^{\frac{n-3}{2}}$ on $[-1,1]$.*

*Proof.* Let the surface area of a unit sphere in $n$ dimensions be $S_n$. We determine the area of a band of sphere from $x_1 = x$ to $x + dx$. If $r = \sqrt{1-x^2}$, the circumference of the band is $S_{n-1}r^{n-2}$, and its width is $\sqrt{dx^2 + \left(\frac{dr}{dx}\,dx\right)^2}$. But $r = \sqrt{1-x^2}$ so $\frac{dr}{dx} = -\frac{x}{\sqrt{1-x^2}}$. Then the width simplifies to

$$\sqrt{1 + \frac{x^2}{1-x^2}}dx = \frac{dx}{\sqrt{1-x^2}},$$

so the total area is

$$\frac{S_{n-1}\sqrt{1-x^2}^{n-2}}{\sqrt{1-x^2}}dx = c(1-x^2)^{\frac{n-3}{2}}dx.$$

Then the probability of picking a point with $a \le X_1 \le b$ is $\int_a^b c(1-x^2)^{\frac{n-3}{2}}dx$ so the density function is $c(1-x^2)^{\frac{n-3}{2}}$ on $[-1,1]$ as desired. □

# B Derivations of gamma function identity

**Lemma 11.** *Up to the constant*

$$c = \frac{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)^2}{2\Gamma\left(\frac{n}{2}\right)^2},$$

*which depends only on $n$,*

$$\sum_{\substack{k=0\\ even}}^{e}\binom{e}{k}\left(\frac{\Gamma\left(\frac{e-k+1}{2}\right)\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{e-k+n}{2}\right)}\right)\left(\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k+2}{2}\right)}\right)\left(\frac{\Gamma\left(e-k+n-1\right)\Gamma\left(k+1\right)}{\Gamma\left(e+n\right)}\right) = c\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{e+1}{2}\right)}{\Gamma\left(\frac{n+e+1}{2}\right)}.$$

*Proof.* The `Mathematica` command

```
Assuming[Element[m,f, Integers],
```

```
FullSimplify[Gamma[(n-1)/2]Gamma[(2f+1)/2]/Gamma[(n+2f+1)/2]/(2Sqrt[Pi])==
Sum[1/Pi Binomial[2f,2j]Gamma[(n)/2]Gamma[(2f-2j+1)/2]
Gamma[(2j+1)/2]Gamma[2m+n-2j-1]Gamma[2j+1]
/(Gamma[(2f+n-2j)/2]Gamma[(2j+2)/2]Gamma[2f+n]),{j,0,f}]]]
```

(where $e$ is replaced by $2f$ and $k$ by $2j$ since each is even) returns `True`, so the identity is verified.

Alternately, for a more human-verifiable proof, we can use the `Mathematica` package `FastZeil`, created by Paule and Schorn [9], the accuracy of which was proven in Schorn [14]. We split into two cases, one where $n$ is even, and one where it is odd. If $n$ is even, we replace it with $2m$ and the command

```
Zb[Binomial[2f,2j]Gamma[(2m)/2]Gamma[(2f-2j+1)/2]
Gamma[(2j+1)/2]Gamma[2f+2m-2j-1]Gamma[2j+1]
/(Gamma[(2f+2m-2j)/2]Gamma[(2j+2)/2]Gamma[2f+2m]), {j,0,f}, m, 1]
```

returns the recursion

```
(1-2m)SUM[m] + (1+2m+2f)SUM[1+m] == 0
```

where `SUM[m]` is the desired sum, which can be checked by hand using the forward difference ($\Delta_k$) also generated by the package:

$$(2m+2f+1)F(j,1+m)+(1-2m)F(j,m)=\Delta_j(F(j,m)R(j,m))$$

where $F(j,m)$ is the desired summand and

$$R(j,m)=\frac{j(-2j+2m+2f-1)}{m+f}.$$

The left side of the equation can be verified to satisfy the recursion. Since the two can be verified to be equal for $m = 1$, they are equal in general.

If $m$ is odd, the analogous approach works, with $m$ replaced by $2m + 1$. The original command becomes

```
Zb[Binomial[2f,2j]Gamma[(2m+1)/2]Gamma[(2f-2l+1)/2]Gamma[(2l+1)/2]
Gamma[2f+2m+1-2j-1]Gamma[2j+1]/Gamma[(2f+2m+1-2j)/2]/
Gamma[(2j+2)/2]/Gamma[2f+2m+1], {j,0,f}, m, 1],
```

and the recursion becomes

```
-m SUM[m] + (1+m+f)SUM[1+m] == 0,
```

with

$$R(j, m) = \frac{2j(-j + m + f)}{2m + 2e + 1}$$

giving

$$(m + f + 1)F(j, 1 + m) - mF(j, m) = \Delta_j(F(j, m)R(j, m)). \quad \square$$

# C Acknowledgments

I would like to thank my mentor, Mr. Gregory Minton, of the Massachusetts Institute of Technology, for all of his guidance and support and my tutor, Dr. John Rickert, of the Rose-Hulman Institute of Technology, for his advice on the writing process. In addition, thanks to Mr. Kartik Venkatram, Dr. Tanya Khovanova, and Prof. David Jerison for matching me with a mentor and guiding me mathematically during the course of my research, and Kate Rudolph, Stella Pantela, and Rohit Agrawal for reading drafts of this paper. Finally, I thank the Massachusetts Institute of Technology, the MIT mathematics department, the Research Science Institute, the Center for Excellence in Education, and Tyco Electronics, without whom none of this would have been possible.