1 Statement of Main Theorem

Any transitive shuffling is asymptotically random.

Definition 1. A shuffling is a probability distribution on $S_n$. That is, $f : S_n \rightarrow \mathbb{R}_{\geq 0}$ such that $\sum_{\pi \in S_n} f(\pi) = 1$.

We think of a shuffling as all the ways we can shuffle a deck of cards, weighted by their probability.

Example 1. Random adjacent transpositions. Or random transpositions (more realistic).

Definition 2. We say a shuffling is transitive if eventually all permutations are possible. That is, there exists a $k$ such that for every $\pi \in S_n$, after $k$ shuffles the probability of the deck being in position $\pi$ is nonzero.

Random transpositions is transitive. Adjacent transpositions has $k = \binom{n}{2}$ and random transpositions has $k = \lceil n/2 \rceil$.

Theorem 1. If $f$ is a transitive shuffling, and $\pi \in S_n$ any permutation, then after shuffling repeatedly, the long-term probability of the deck landing in permutation $\pi$ is $\frac{1}{n!}$.

Any transitive shuffling will eventually randomize the deck.

2 History: Poincaré and Persi Diaconis

- 1923 Poincaré, Henri Calcul des probabilités. 2nd Ed.
  - Appears as a theorem in Poincaré’s only probability textbook. (12pg proof)

- 1992 Card shuffling popularized by Persi Diaconis and Dave Bayer’s paper Trailing the dovetail shuffle to its lair. They showed 7 riffle shuffles is sufficient to randomize a deck of 52 cards.
• Persi Diaconis (Wikipedia bio)
  – Born January 31, 1945
  – Left home at 14 to travel with sleight-of-hand legend Dai Vernon, and dropped out of high school, promising himself that he would return one day so that he could learn all of the math necessary to read William Feller’s famous two-volume treatise on probability theory, *An Introduction to Probability Theory and Its Applications*. He returned to school (City College of New York for his undergraduate work graduating in 1971 and then a Ph.D. in Mathematical Statistics from Harvard University in 1974), learned to read Feller, and became a mathematical probabilist.
  – According to Martin Gardner, at school, Diaconis supported himself by playing poker on ships between New York and South America. Gardner recalls that Diaconis had "fantastic second deal and bottom deal".
  – Known for studying card shuffling, coin flipping, and dice rolling.

3 Group Algebra statement

Recall the *group algebra* $\mathbb{R}[G]$ of a finite group $G$, the set of all functions $f : G \rightarrow \mathbb{R}$ endowed with scalar multiplication, pointwise $+$, and $\ast$ convolution. It’s the same as formal sums, of course.

**Example 2.** Suppose $f$ is uniform on adjacent transpositions. Then the powers of $f$ are:

<table>
<thead>
<tr>
<th>n</th>
<th>(12)</th>
<th>(13)</th>
<th>(23)</th>
<th>(123)</th>
<th>(132)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>19</td>
<td>16</td>
<td>8</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>51</td>
<td>46</td>
<td>30</td>
<td>46</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>143</td>
<td>132</td>
<td>100</td>
<td>132</td>
<td>111</td>
</tr>
<tr>
<td>7</td>
<td>407</td>
<td>386</td>
<td>322</td>
<td>386</td>
<td>343</td>
</tr>
<tr>
<td>8</td>
<td>1179</td>
<td>1316</td>
<td>1008</td>
<td>1136</td>
<td>1051</td>
</tr>
<tr>
<td>9</td>
<td>3451</td>
<td>3366</td>
<td>3110</td>
<td>3366</td>
<td>3195</td>
</tr>
<tr>
<td>10</td>
<td>10183</td>
<td>10012</td>
<td>9500</td>
<td>10012</td>
<td>9671</td>
</tr>
<tr>
<td>11</td>
<td>30207</td>
<td>29866</td>
<td>28842</td>
<td>29866</td>
<td>29183</td>
</tr>
<tr>
<td>12</td>
<td>89939</td>
<td>89256</td>
<td>87208</td>
<td>89256</td>
<td>87891</td>
</tr>
</tbody>
</table>

**Remark 1.** For $f$ defined above, normalizing gives a shuffling on $S_3$. Squaring $f$ corresponds to shuffling twice.

**Example 3.** Replacing $f$ with $\frac{1}{3}f$ the powers of $f$ are:
In the Group Algebra, our Main Theorem becomes:

**Theorem 2.** Let \( f \in S_n \) be a transitive shuffling. Then for any \( \pi \in S_n \),

\[
\lim_{s \to \infty} f^s(\pi) = \frac{1}{n!}
\]  

Remark 2. Poincaré also used the group algebra in his proof, calling them “complex numbers”. Poincaré’s approach was to compute the eigenvalues and use them to describe the asymptotic growth.

Remark 3. Feller’s famous textbook *An Introduction to Probability Theory and Its Applications* Sec. 15.10 (Markov Chains), also mentions this theorem. Feller derives the result by standard Markov Chain technology (the transition matrix is doubly stochastic).

## 4 Lemmas and proof

Let \( G \) be an arbitrary finite group. If \( f \in \mathbb{R}[G] \), \( \int f := \sum_{g \in G} f(g) \). Note that

- \( \int \alpha f = \alpha \int f \)
- \( \int f + g = \int f + \int g \)
- \( \int (f * g) = (\int f)(\int g) \) (haha, surprised? Note: \( \int \) is just \( \rho_{\text{triv}} \))

Let \( \mathbb{R}_P[G] := \{ f \in \mathbb{R}[G] \mid \int f = 1 \} \). Closed under convolution.

Our proof requires one simple Lemma:

**Lemma 1.** Let \( U \) be the constant function on \( G \) with value \( \frac{1}{|G|} \), and let \( f \in \mathbb{R}_P[G] \) be any other distribution. Then \( U * f = f * U = U \).

**Proof.** It’s easy to prove \( U * g = U \) for any \( g \in G \). By linearity, it follows that \( U * f = \alpha U \). Since \( U * f \in \mathbb{R}_P[G] \), we have \( \alpha = 1 \).

**Corollary 1.** The \( U \) defined above is an idempotent, that is \( U^k = U \) for any \( k \geq 1 \).
4.1 Proof of Main Theorem

Since \( f \) is transitive, \( f^k = pT + qU \) where \( U \) is as above, and \( T \) is the normalized “remainder”, with \( \int T = 1 \), so \( p + q = 1 \). Now take a large power of \( pT + qU \) and use the Binomial Theorem\(^1\):

\[
(f^k)^N = (pT + qU)^N = \sum_{i=0}^{N} \binom{N}{i} (pT)^i(qU)^{N-i}
\]

\[= \sum_{i=0}^{N-1} \binom{N}{i} p^i q^{N-i} U^{N-i} + p^N T^N \tag{4.1}
\]

\[= U \sum_{i=0}^{N-1} \binom{N}{i} p^i q^{N-i} + p^N T^N \tag{4.2}
\]

\[= (1 - p^N)U + p^N T^N \tag{4.3}
\]

Since \( p < 1 \), as \( N \to \infty \), only \( U \) survives. This shows that \( \lim_{N \to \infty} (f^k)^N = U \). It follows (with a step or two more\(^2\)) that \( \lim_{N \to \infty} f^N = U \).

5 How fast to random?

Our argument shows that the convergence is eventually geometrically fast. But some shuffling schemes approach uniform faster than others. Practically speaking, this is an important question.

5.1 Random to Top (Tsetlin Library)

By the coupon collector problem, Random to Top becomes randomized after approximately \( n \ln n \) operations. Application to “recently used files”. Note for 52 that’s about 200.

5.2 Random Adjacent Transpositions

It was shown that \( n^3 \ln n \) is sufficient (David Aldous 1982). The bound was improved to \( \frac{n^2}{\pi^2} \ln n \) by David Wilson in 2002.

5.3 Random Transpositions

Diaconis and Shahshahani (1981) were able to show that the mixing time is approximately \( \frac{1}{2} n \ln n \). The paper is called Generating a random permutation with

---

\(^1\)The Binomial Theorem \( (a + b)^N = \sum_{i=0}^{N} \binom{N}{i} a^i b^{N-i} \) is true in any ring provided that \( ab = ba \) (not true for elements that don’t commute). In this case, \( U \) is a central element, and so the theorem applies.

\(^2\)Easy to show from (4.4) that multiplying by more \( f \)’s maintains the \( (1 - p^N)U \) part, so can’t bring you further from \( U \) than \( p^N \). Assuming \( 0 \leq f(g) \leq 1 \).
random transpositions. The techniques are representation theory of $S_n$ (standard young tableaux, etc.). This is because the group algebra is decomposable into the irreducible representations.

5.4 Dovetail/riffle shuffle

In 1992, Persi Diaconis and David Bayer famously proved that 7 dovetail shuffles are sufficient (in general $\approx \frac{3}{2} \log_2 n$). The method does not use representation theory, but just combinatorics (in fact, Eulerian numbers!).

6 Connection to Central Limit Theorem

The result fails for infinite groups, as soon as $G = \mathbb{Z}$. In fact, the Central Limit Theorem answers the question there: $f \in \mathbb{R}[\mathbb{Z}]$ will converge to a bell curve with parameters $E[f]$ and $SD(f)$. Pretty amazing (although remember the bell curve is the only shape stable under convolution, which is an analog to our theorem about the idempotent $U$). As a corollary of these results, the bell curve mod $p$ is uniform (is that surprising?).

7 Application to $\mathbb{Z}_p$ and RNGs

If you need to generate a random integer, maybe you can take some small source of randomness from nature and then convolve several times in $G = \mathbb{Z}_p$ for some $p$. You don’t really need to know “how random” the nature is, our theorem shows that if you just convolve enough times you get a uniformly random number. This is (sort of) how pseudorandom number generators work.

8 Tits Algebra factoid

I was shocked to stumble upon a paper by Persi Diaconis while I was researching hyperplane arrangements. It turns out that the faces of a hyhperplane arrangement have a certain structure known as the Tits algebra, which is still not well-understood for many arrangements. Diaconis and Brown (1997) proved that for the braid arrangement this structure is equivalent to riffle shuffling. The result was expanded upon by Diaconis and Christos Athanasiadis (!) in 2010. Awesome.