3. WEYL'S EQUIDISTRIBUTION THEOREM

Let $x \in [0,1)$ and consider the sequence $\overline{nx} := nx \pmod 1$. What can we say about this sequence? If $x = p/q$ is a rational number, then in $\overline{(qm+r)x} = \overline{rx}$, so what we have is a finite sequence that repeats itself endlessly. What is x is an irrational number? It is clear that $\overline{nx}$ are all distinct. It is also easy to show that the sequence is dense in $[0,1]$.

Exercise 1. If $x \in [0,1] \setminus \mathbb{Q}$, show that $\{\overline{nx} : n \geq 1\}$ is dense in $[0,1]$.

What we want to know is the amount of 'time' the sequence spends in various subsets of $[0,1]$. For example, does it spend equal time in $[0.10, 0.25]$ as in $[0.5, 0.65]$? Observe that working modulo 1 is equivalent to considering the circle group. In other words, we may consider the sequence $e^{2\pi ikx}$ and ask about the proportion of 'time' it spends in a given arc of the unit circle.

Theorem 2 (Weyl). *Let $x \in [0,1] \setminus \mathbb{Q}$. Then, for any interval $I \subseteq [0,1]$, we have*

$$\frac{1}{n}\#\{k \leq n : \overline{kx} \in I\} \to |I| \quad \text{as } n \to \infty.$$

More generally, one may ask the same question for any sequence.

Definition 3. A sequence $(a_n)_n$ of real numbers is said to be equidistributed in $[0,1]$ if

$$\frac{1}{N}\#\{k \leq N : \overline{a_k} \in I\} \to |I|, \quad \text{as } N \to \infty \tag{1}$$

for every interval $I \subseteq [0,1]$.

In this language, Theorem 2 says that $(n\alpha)_n$ is equidistributed if α is irrational. What about the sequence $(n^2\alpha)_n$ or $(n^3\alpha)_n$ etc? Are they also equidistributed? We shall show the following result.

Theorem 4 (Weyl). *Let $P(x) = x^2 + bx + c$ be a monic quadratic polynomial with real coefficients. If α is irrational, then the sequence $(\alpha P(n))_n$ is equidistributed. In particular, the sequence $(\alpha n^2)_n$ is equidistributed.*

In fact, Weyl proved the amazing result that if P is any polynomial in which at least one of the coefficients other than the constant coefficient is irrational, then $(P(n))_n$ is equidistributed. We shall not prove this result, but the proof is only a small extension of the proof we give for Theorem 4.

One might be tempted to think that if $f : \mathbb{R}_+ \to \mathbb{R}$ is any function which has something "irrational" about it, then the sequence $(f(n))_n$ becomes equidistributed. This is part of a larger fallacy that if we cannot see a pattern then it must be truly random! Here is a counterexample.

Example 5. Let $\alpha = \frac{\sqrt{5}+1}{2}$. Then the sequence α^n is not equidistributed. Indeed, if $\beta = \frac{1-\sqrt{5}}{2}$, then α and β are the two roots of $x^2 - x - 1$. Then of $\alpha + \beta = 1$ and $\alpha\beta = -1$, from which it follows (how?) that $\alpha^n + \beta^n$ is an integer for all n. Thus, $\overline{\alpha^n}$ is either β^n or $1 - \beta^n$. But $\beta^n \to 0$ and hence $(\overline{\alpha^n})_n$ is not even dense (it has only two limit points, 0 and 1).

Here is another simple example.

Exercise 6. The sequence $(n!e)_n$ is not equidistributed.

We now set up the basic technique for proving equidistribution. Theorem 2 will be an immediate consequence, whereas Theorem 4 will need one more technique.

The basic technique - exponential sums: In terms of the indicator function $\mathbf{1}_I$ of the interval, we may write the left hand side of as $\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_I(\overline{kx})$. That naturally suggests a generalization to functions by replacing $\mathbf{1}_I$ by another function. The key to the proof is to prove the statement for continuous functions and even there, prove it only for a few special continuous functions. Here and later, let $e(x) := e^{2\pi ix}$.

Lemma 7. *Let $f \in C(S^1)$, i.e., $f : [0,1] \to \mathbb{R}$ be continuous and satisfy $f(0) = f(1)$. Then, $(a_n)_n$ is equidistributed if and only if*

$$\frac{1}{n}\sum_{k=1}^{n} e(\ell a_k) \to 0 \quad \textit{for all } \ell \in \mathbb{Z}\setminus\{0\}. \tag{2}$$

We shall only show that (2) implies (1) as that is what is needed to show equidistribution. The other way implication is a homework problem.

Proof that (2) *implies* (1). Suppose (2) holds. Observe that when $\ell = 0$, the left hand side of (2) converges to 1. We now claim that for any $f \in C[0,1]$ with $f(0) = f(1)$ (i.e., $f \in C(S^1)$), we have

$$\frac{1}{n}\sum_{k=1}^{n} f(\overline{a_k}) \to \int_0^1 f(t)dt. \tag{3}$$

Indeed, (2) implies the above convergence for all trigonometric polynomials. By Fejér's theorem, given f and ε there is a trigonometric polynomial T such that $\|f - T\|_{\sup} < \varepsilon$. Then both the left hand and right hand sides of (3) are within ε of the corresponding quantities when f is replaced by T. The rest is obvious.

Now we prove equidistribution. Let $I = [a,b]$. Fix $\varepsilon > 0$ and choose two continuous functions $f, g \in C(I)$ such that $\mathbf{1}_{[a+\varepsilon, b-\varepsilon]} \le g \le \mathbf{1}_I \le f \le \mathbf{1}_{[a-\varepsilon, b+\varepsilon]}$ (note that everything here is modulo 1. If $a = 0$ then $a - \varepsilon = 1 - \varepsilon$ etc). Then,

$$\int_0^1 g(t)dt = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} g(\overline{a_k}) \le \liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_I(\overline{a_k}) \le \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_I(\overline{a_k}) \le \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(\overline{a_k}) = \int_0^1 f(t)dt.$$

Also, $\int_0^1 f - \int_0^1 g \le 4\varepsilon$. As ε is arbitrary, the implication follows. ■

Remark 8. Approximating by an easier class of functions is one of the basic techniques in proofs in analysis. The sense of approximation should always be adapted to the need at hand. For instance, approximation was used twice in the above proof - once when approximating continuous functions by trigonometric polynomials (sense of approximation: uniform) and second time when approximating indicator of an interval by a continuous function (by sandwiching).

All proofs we give of equidistribution will be by proving (2). It works like a charm in the first theorem.

Proof of Theorem 2. Fix $\ell \in \mathbb{Z}\setminus\{0\}$. Then,

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n} e(\ell kx) &= \frac{1}{n}\sum_{k=1}^{n} e^{2\pi i \ell kx} \\ &= \frac{1}{n} e^{2\pi i \ell x} \frac{e^{2\pi i \ell nx} - 1}{e^{2\pi i \ell x} - 1}.\end{aligned}$$

Since x is irrational, $e^{2\pi i \ell x} \neq 1$, which validates the above expression. Further, the last term is bounded in absolute value by $\frac{2}{n|e^{2\pi i \ell x} - 1|}$ which goes to zero as $n \to \infty$. Since $\int_I e^{2\pi i m x} dx = 0$ for $\neq 0$, we have proved (2). By Lemma 7 it follows that $(kx)_k$ is equidistributed. ■

The following quantitative bound will be useful later. In this section, $\|x\| := \min\{|x - n| : n \in \mathbb{Z}\}$ is the distance from x to the nearest integer (that is, to $\lfloor x \rfloor$ or $\lceil x \rceil$, whichever is closer). Further, $x \wedge y$ means $\min\{x, y\}$.

Exercise 9. For any $\alpha \in \mathbb{R}$ show the bound $\left|\sum_{k=1}^{n} e(k\alpha)\right| \le \frac{1}{2\|\alpha\|} \wedge n$.

The same proof will not work when $a_k = k^2$ because we do not have an explicit form for the sum $\sum_{k=1}^{n} e^{2\pi i \ell k^2 x}$. We shall need one more important technique given in the following lemma.

Lemma 10. *[Weyl's differencing method] Let $f : [n] \to \mathbb{C}$ be any function. Then,*

$$\Big|\sum_{k=1}^{n} e(f(k))\Big|^2 \le 2\sum_{m=0}^{n-1}\Big|\sum_{j=1}^{n-m} e(f(m+j)-f(j))\Big|$$

Proof. Since $\overline{e(x)} = e(-x)$, we write

$$\begin{aligned}\Big|\sum_{k=1}^{n} e(f(k))\Big|^2 &= \sum_{k,\ell=1}^{n} e(f(k)-f(\ell)) \\ &= n + \sum_{m=1}^{n-1}\sum_{j=1}^{n-m} e(f(j+m)-f(j)) + \sum_{m=1}^{n-1} e(f(j)-f(j+m)) \\ &= n + 2\sum_{m=1}^{n-1} \mathrm{Re}\left\{\sum_{j=1}^{n-m} e(f(j+m)-f(j))\right\} \\ &\le n + 2\sum_{m=1}^{n-1}\Big|\sum_{j=1}^{n-m} e(f(j+m)-f(j))\Big|.\end{aligned}$$

This last quantity is clearly bounded by the right hand side in the statement of the theorem. ■

Why is this useful? The idea is that when f is a polynomial of degree 2, then for each m, the difference $x \to f(x+m) - f(x)$ is a polynomial of degree 1. As equidistribution for αn has been proved, each of the inner sums on the right hand side are small when n is large. We can hope to prove that the left hand side is also small. However, the outer sum has an increasing number of terms (i.e., depends on n too). Therefore, more care is needed in carrying out the details.

Note: I did not get around to writing the complete notes for this part. Some of you have the notes from my lecture. If you bring it to me, we can scan it and put it online