

Weyl's equidistribution theorem for linear and quadratic polynomials

Sayantana Khan

July 2016

Given a polynomial $p \in \mathbb{R}[X]$, one can look at the orbit of the polynomial restricted to the interval $[0, 1]$. The orbit is the set

$$\{p(n) \bmod 1 \mid n \in \mathbb{N}\}$$

If $p(x)$ is the polynomial ax , where a is an irrational number, then it's not very difficult to show that the orbit is dense in $[0, 1]$. Conversely, if a is rational, then the orbit is not dense. We can actually expect something stronger than density, namely equidistribution. Informally, a sequence is equidistributed if the fraction of times it appears in any subinterval is proportional to the measure of the subinterval. We'll show that if p is a polynomial of degree 1 or 2 with an irrational leading coefficient, then the orbit is equidistributed.

1 Equidistributed sequences

We'll begin by formally defining what equidistributed sequences are.

Definition 1.1. A sequence $\{q_n\}$ is said to be equidistributed in the interval $[a, b]$ if

- $q_n \in [a, b]$ for all $n \in \mathbb{N}$.
- For all subintervals $[c, d]$ of $[a, b]$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \mathbb{1}_{[c,d]}(q_k) = \frac{d-c}{b-a}$$

where $\mathbb{1}_{[c,d]}$ is the indicator function for the interval $[c, d]$.

For now, let's just focus on the interval $[0, 2\pi]$, and look for equidistributed sequences in it. Observe that the second part of the above definition can be rewritten as

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \mathbb{1}_{[c,d]}(q_k) = \frac{1}{2\pi} \int_0^{2\pi} \mathbb{1}_{[c,d]}(x) dx \quad (1)$$

2 Weyl's criterion for equidistribution

The above definition for equidistributed sequences makes sense (in the manner that agrees with our intuition), it is a little cumbersome to work with, since one needs to sum the indicator function of every subinterval of $[0, 2\pi]$ over the entire sequence to ensure it really is equidistributed. It would be nicer if one had a simpler, but equivalent notion for equidistribution that was easier to work with. It turns out there is one.

Lemma 2.1. *A subsequence $\{q_k\}$ of $[0, 2\pi]$ is equidistributed iff for every continuous function f on $[0, 2\pi]$*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N f(q_k) = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx$$

Proof. First we'll show that if the sequence is equidistributed then the above equation is correct. Consider *simple functions*¹ on $[0, 2\pi]$. Since these functions are linear combinations of indicator functions, and the indicator functions satisfy equation 1, the simple functions also satisfy that equation, i.e.

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N s(q_k) = \frac{1}{2\pi} \int_0^{2\pi} s(x) dx$$

where s is a simple function. Also, note the fact that in the compact interval $[0, 2\pi]$, for every $\varepsilon > 0$ and for every continuous function f , there exists a simple function s , such that $\|f - s\|_\infty < \varepsilon$.

Now we'll show that for every $\varepsilon > 0$, and for every continuous function f , there exists an $N \in \mathbb{N}$ such that for all $n > N$

$$\left| \frac{1}{n} \sum_{k=1}^n f(q_k) - \frac{1}{2\pi} \int_0^{2\pi} f(x) dx \right| < \varepsilon$$

Pick a simple function s such that $\|f - s\|_\infty < \frac{\varepsilon}{3}$. We'll have the following inequalities

$$\frac{1}{2\pi} \left| \int_0^{2\pi} f(x) dx - \int_0^{2\pi} s(x) dx \right| \leq \frac{1}{2\pi} \left| \int_0^{2\pi} (f(x) - s(x)) dx \right| \quad (2)$$

$$\leq \frac{\varepsilon}{3} \quad (3)$$

And for all $n \in \mathbb{N}$

$$\left| \frac{1}{n} \sum_{k=1}^n f(q_k) - \frac{1}{n} \sum_{k=1}^n s(q_k) \right| \leq \left| \frac{1}{n} \sum_{k=1}^n f(q_k) - s(q_k) \right| \quad (4)$$

$$\leq \frac{\varepsilon}{3} \quad (5)$$

¹Simple functions are functions whose range is finite. These functions can be written as linear combinations of indicator functions.

Now pick an N large enough such that for all $n > N$

$$\left| \frac{1}{n} \sum_{k=1}^n s(q_k) - \frac{1}{2\pi} \int_0^{2\pi} s(x) dx \right| < \frac{\varepsilon}{3} \quad (6)$$

Adding up inequalities 3, 5, and 6, we have the inequality we wanted.

Now for the converse statement. The key idea here is to approximate the indicator function $\mathbb{1}_{[c,d]}$ with an appropriate continuous function. For a given $\varepsilon > 0$, consider the following continuous approximation for the function $\mathbb{1}_{[c,d]}$

$$f(x) = \begin{cases} 1 & x \in (c, d) \\ \frac{8x-8c+\varepsilon}{\varepsilon} & x \in [c - \frac{\varepsilon}{8}, c] \\ \frac{8d-8x+\varepsilon}{\varepsilon} & x \in [d, d + \frac{\varepsilon}{8}] \\ 0 & \text{otherwise} \end{cases}$$

For all $x \in [0, 2\pi]$, we have the following inequality for the function f

$$\mathbb{1}_{[c,d]}(x) \leq f(x) \leq \mathbb{1}_{[c-\frac{\varepsilon}{8}, d+\frac{\varepsilon}{8}]}(x)$$

Rewriting the above inequality,

$$|f(x) - \mathbb{1}_{[c,d]}(x)| \leq \mathbb{1}_{[c-\frac{\varepsilon}{8}, c]}(x) + \mathbb{1}_{[d, d+\frac{\varepsilon}{8}]}(x)$$

This means there exists a large enough N_1 , such that for all $n > N_1$

$$\frac{1}{n} \sum_{k=1}^n \left(\mathbb{1}_{[c-\frac{\varepsilon}{8}, c]}(q_k) + \mathbb{1}_{[d, d+\frac{\varepsilon}{8}]}(q_k) \right) \leq \frac{\varepsilon}{3}$$

For that same N_1 ,

$$\left| \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{[c,d]}(q_k) - \frac{1}{n} \sum_{k=1}^n f(q_k) \right| \leq \frac{\varepsilon}{3} \quad (7)$$

Similarly,

$$\frac{1}{2\pi} \left| \int_0^{2\pi} f(x) dx - \int_0^{2\pi} \mathbb{1}_{[c,d]}(x) dx \right| \leq \frac{\varepsilon}{3} \quad (8)$$

Finally, because of our hypothesis, we have a large enough N_2 such that for all $n > N_2$,

$$\left| \frac{1}{n} \sum_{k=1}^n f(q_k) - \frac{1}{2\pi} \int_0^{2\pi} f(x) dx \right| \leq \frac{\varepsilon}{3} \quad (9)$$

Let $N = \max(N_1, N_2)$, and adding up inequalities 7, 8, and 9, we get our result.

This completes the proof of equivalence. \square

This result shows how to equidistribution in terms of continuous functions. The next result will use Fejér's theorem to break up the result about continuous functions to exponential sums, giving us a much simpler formalism to work with.

Theorem 2.2. A subsequence $\{q_k\}$ of $[0, 2\pi]$ is equidistributed iff for all non-zero integers a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \exp(iaq_k) = 0$$

Proof. First we'll show that if the sequence is equidistributed, then the limit is 0. Since $\exp(iax)$ is a continuous function for all a , from lemma 2.1, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \exp(iaq_k) &= \frac{1}{2\pi} \int_0^{2\pi} \exp(iax) dx \\ &= 0 \end{aligned} \quad (a \text{ is non-zero})$$

For the converse, we will use Fejér's theorem. Given a continuous f and $\varepsilon > 0$, use Fejér's theorem to find an exponential polynomial p such that $\|p - f\|_\infty < \frac{\varepsilon}{3}$. In that case, we have

$$\frac{1}{2\pi} \left| \int_0^{2\pi} f(x) dx - \int_0^{2\pi} p(x) dx \right| < \frac{\varepsilon}{3} \quad (10)$$

and

$$\left| \frac{1}{n} \sum_{k=1}^n f(q_k) - \frac{1}{n} \sum_{k=1}^n p(q_k) \right| < \frac{\varepsilon}{3} \quad (11)$$

And because of our hypothesis, we have a large enough N , such that for all $n > N$

$$\left| \frac{1}{n} \sum_{k=1}^n p(q_k) - \int_0^{2\pi} p(x) dx \right| \leq \frac{\varepsilon}{3} \quad (12)$$

Adding up inequalities 10, 11, and 12, we get our result. \square

3 Weyl's equidistribution theorem: The weak version

The weak version of Weyl's equidistribution theorem states that the sequence

$$q_k = kz \pmod{2\pi}$$

is equidistributed in the interval $[0, 2\pi]$ if z is an irrational multiple of π . This can be shown really easily using the earlier formalisation, by looking at the exponential sums.

$$\sum_{k=1}^n \exp(iakz) = e^{iaz} \frac{1 - e^{ianz}}{1 - e^{iaz}}$$

Since z is an irrational multiple of π , the denominator is non-zero for all non-zero integers a . That means for a given a , the sum is bounded, hence

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \exp(iakz) = 0$$

and as a consequence, the sequence is equidistributed.

4 Weyl's equidistribution theorem: The stronger version

Note: Since it will get a little cumbersome to keep working with integral and irrational multiples of 2π , we will instead deal with sequences in the interval $[0, 1]$. The only change we'll need to keep in mind is that we'll be dealing the following sum

$$\sum_{k=1}^n \exp(2\pi i a q_k)$$

The stronger version of Weyl's equidistribution theorem states that the sequence

$$q_k = \alpha k^2 + \beta k + \gamma \pmod{1}$$

is equidistributed if α is irrational.

This will involve showing that the sequence

$$S_n = \sum_{k=1}^n \exp(2\pi i a q_k) \tag{13}$$

is $o(n)$ for all integers n , but since α is irrational, it suffices to show it for all irrational α . In a nutshell, we have reduced the problem of equidistribution to a problem of bounding the given sum.

4.1 A preliminary bound for another sum

Notation: We will define the function $e(x)$ to be $e(2\pi i x)$.

Lemma 4.1. *For all irrational θ*

$$\left| \sum_{k=1}^N e(n\theta) \right| \leq \min \left(N, \frac{1}{\|\theta\|} \right) \tag{14}$$

where $\|\theta\|$ is distance of θ to its closest integer.

Proof. By triangle inequality,

$$\begin{aligned} \left| \sum_{k=1}^N e(n\theta) \right| &\leq \sum_{k=1}^N |e(n\theta)| \\ &= N \end{aligned}$$

And summing up the geometric series, we have

$$\begin{aligned} \left| \sum_{k=1}^N e(n\theta) \right| &\leq \frac{2}{|1 - e(\theta)|} \\ &= \frac{1}{|\sin(\pi\theta)|} \\ &\leq \frac{1}{\|\theta\|} \quad \square \end{aligned}$$

4.2 Weyl differencing

Since we want to show the sum S_N (see equation 13) is in $o(N)$, it will suffice to show that $|S_N|^2$ is in $o(N^2)$.

$$\begin{aligned}
\left| \sum_{k=1}^N e(f(k)) \right|^2 &= \left(\sum_{k=1}^N e(f(k)) \right) \left(\sum_{k=1}^N e(-f(k)) \right) \\
&= \sum_{1 \leq j, l \leq N} e(f(j) - f(l)) \\
&= N + \sum_{d=1}^{N-1} \sum_{k=1}^{N-d} (e(f(k+d) - f(k)) + e(f(k) - f(k+d))) \\
&= N + \sum_{d=1}^{N-1} \sum_{k=1}^{N-d} 2\operatorname{Re}(e(f(k+d) - f(k))) \\
&= N + 2 \sum_{d=1}^{N-1} \operatorname{Re} \left(\sum_{k=1}^{N-d} e(f(k+d) - f(k)) \right) \\
&\leq N + 2 \sum_{d=1}^{N-1} \left| \sum_{k=1}^{N-d} e(f(k+d) - f(k)) \right|
\end{aligned}$$

Plugging in the quadratic polynomial in f , we get

$$\begin{aligned}
\left| \sum_{k=1}^N e(q_k) \right|^2 &\leq N + 2 \sum_{d=1}^{N-1} \left| \sum_{k=1}^{N-d} e((2\alpha d)k) \right| \\
&\leq N + 2 \sum_{d=1}^N \min \left(N, \frac{1}{\|2\alpha d\|} \right)
\end{aligned}$$

Since N is in $o(N^2)$, all we need to show is that $\sum_{d=1}^N \min \left(N, \frac{1}{\|2\alpha d\|} \right)$ is also in $o(N^2)$.

4.3 Rational approximation of irrational numbers

The main result of this section will show that given an irrational number α , it's possible to find infinitely many rational numbers $\frac{p}{q}$ ($\gcd(p, q) = 1$) such that $|\alpha - \frac{p}{q}| < \frac{1}{q^2}$. We will need this result to bound the main sum.

Proposition 4.2. *Given an irrational number α and $N \in \mathbb{N}$, there exists a rational number $\frac{p}{q}$ such that*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{q(N+1)}$$

and $|q| \leq N$.

Proof. Consider the following sequence modulo 1:

$$\alpha, 2\alpha, 3\alpha, \dots, N\alpha$$

Divide $[0, 1]$ into $N + 1$ equally sized intervals. Clearly, one of the elements of that sequence must either lie in the subinterval $[0, \frac{1}{N+1}]$, or $[\frac{N}{N+1}, 1]$. Call that element $q\alpha$. In that case, the following inequality holds:

$$\|q\alpha\| \leq \frac{1}{N+1}$$

Let the integer closest to $q\alpha$ be p . In that case

$$\begin{aligned} |p - q\alpha| &\leq \frac{1}{N+1} \\ \left| \frac{p}{q} - \alpha \right| &\leq \frac{1}{q(N+1)} \end{aligned}$$

And since $q < N + 1$

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{q^2} \quad \square$$

Corollary 4.3. *If α is irrational, there are infinitely many rational numbers $\frac{p}{q}$ such that*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{q^2}$$

Proof. Suppose there were finitely many. Take the minimal distance d of α with the rational approximations. Pick a large enough N such that $\frac{1}{N+1} < d$. Use that N in proposition 4.2 and get a contradiction. \square

4.4 Bounding the complete sum[1]

At the end of section 4.2, we saw that all we need to do is to show that the following sum

$$W_N = \sum_{d=1}^N \min \left(N, \frac{1}{\|2\alpha d\|} \right)$$

is in $o(N^2)$.

Lemma 4.4. *If $\alpha_1, \alpha_2, \dots, \alpha_N$ are real numbers such that $\|\alpha_i - \alpha_j\| \geq \frac{1}{r}$ for some natural number r when $i \neq j$. Then*

$$\sum_{i=1}^N \min \left(\frac{1}{\|\alpha_i\|}, N \right) \leq 2N + 2r(\log(N) + 2)$$

Proof. Without loss of generality, assume all the α_i lie within $[-\frac{1}{2}, \frac{1}{2}]$. Furthermore, at least half the sum will either come from the positive or the negative α_i . WLOG, assume it's the positive α_i . Then we can just double the sum to get an upper bound. Assume that $\alpha_1 < \alpha_2 < \dots < \alpha_n$, where α_1 to α_n are the positive α_i . In that case

$$\sum_{i=1}^n \min\left(\frac{1}{\|\alpha_i\|}, N\right) \leq N + \left\lfloor \frac{r}{N} \right\rfloor N + \sum_{i > \lfloor \frac{r}{N} \rfloor} \frac{r}{i} \quad \left(\text{Because } \|\alpha_i - \alpha_j\| > \frac{1}{r}\right)$$

The last term can have at most n terms, hence it is bounded by $r(\log(n) + 1)$. But n is smaller than or equal to N , since n is the number of positive α_i . Hence the last term is bounded by $r(\log(N) + 1)$. The second term is bounded by r , and the first term is N . Hence the complete bound is

$$\sum_{i=1}^n \min\left(\frac{1}{\|\alpha_i\|}, N\right) \leq N + r + r(\log(n) + 1)$$

Since this is at least half the total sum, we get our bound. \square

Lemma 4.5. *If $\frac{p}{q}$ is a rational approximation for α such that $\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}$, then*

$$\sum_{d=1}^N \min\left(\frac{1}{\|\alpha d\|}, N\right) \leq \left(\frac{2N}{q}\right) (2N + 4q(\log(N) + 2))$$

Proof. We have the following inequality, which is not too hard to prove:

$$\|i\alpha - j\alpha\| \geq \left\| (i-j)\frac{p}{q} \right\| - \frac{|i-j|}{q^2} \quad \left(\text{Taylor expand around } \frac{p}{q}\right)$$

When $|i-j| \leq \frac{q}{2}$, then $\frac{(i-j)p}{q}$ is not an integer since p and q are co-prime. That means

$$\left\| (i-j)\frac{p}{q} \right\| \geq \frac{1}{q}$$

and

$$\frac{|i-j|}{q^2} \leq \frac{1}{2q}$$

which means

$$\|i\alpha - j\alpha\| \geq \frac{1}{2q}$$

Now we split up the range of summation, i.e. the interval $[1, N]$ into intervals of length $\frac{q}{2}$. To compute the sum over each of these intervals, we'll use lemma 4.4 with $r = 2q$. We'll have $\frac{2N}{q}$ such intervals, over each of which, the sum will be bounded by $(2N + 4q(\log(N) + 2))$. The complete bound is hence

$$W_N \leq \left(\frac{2N}{q}\right) (2N + 4q(\log(N) + 2)) \quad \square$$

As N goes to infinity, the $\frac{W_N}{N^2}$ becomes $\frac{2}{q}$. But since α is irrational, q can be as large as possible. This shows that W_N is in $o(N^2)$.

References

- [1] Jacques Verstraëte, *Notes on equidistribution*, <http://www.math.ucsd.edu/~jverstra/Weyl2.pdf>.