Approximation of irrationals by rationals

Elias Dubno

March 12th, 2020

1 Convergents as best approximations

Recall the following definition.

Definition 1.1. If $x = [a_0; a_1, a_2, \dots]$ is a continued fraction of x, then the rational number

$$\frac{p_n}{q_n} = [a_0; a_1, \dots, a_n] \qquad (n \geq 0)$$

is called the n-th *convergent* to x.

Example 1.2. Consider $x = \pi$. We have seen in the first talk that the first convergent to π is $\frac{p_1}{q_1} = \frac{22}{7} \approx 3.142$. We notice that two digits after the decimal point are already correct. In order to get this level of precision by cutting of the decimal expansion, we would have $3.14 = \frac{314}{100} = \frac{157}{50}$. Note that the numbers 22 and 7 are much smaller than 157 and 50, but give almost the same precision. This is an indicator that convergents are very good approximations.

Our goal is not only to show that the convergents are very good approximations, but in fact the best approximations. But first, we need to define the notion of *best approximation*.

Definition 1.3. Let $x \in \mathbb{R}$. We define *the distance of x to the nearest integer* by

$$\|x\| = \min\{|n - x| : n \in \mathbb{Z}\}.$$

Remark 1.4. The distance to the nearest integer $\|\cdot\|$ does not define a norm, since it is not homogeneous. For example:

$$\left\|\frac{1}{2} \cdot \frac{2}{3}\right\| = \left\|\frac{1}{3}\right\| = \frac{1}{3} \neq \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = \left|\frac{1}{2}\right| \cdot \left\|\frac{1}{3}\right\|.$$

Definition 1.5. Let $x \in \mathbb{R}$. An integer $q > 0$ is called a *best approximation* to x if

$$\forall q' \in \mathbb{N} : q' < q \implies \|q'a\| > \|qa\|.$$

We can now state the main Theorem of this section.

Theorem 1.6. *Let $x \in \mathbb{R} \setminus \mathbb{Q}$. An integer $q > 0$ is a best approximation to x if and only if $q = q_n$ (the denominator of the n-th convergent) for some $n \geq 1$.*

In order to prove Theorem 1.6, we need the following Lemma:

Lemma 1.7. *Let $\frac{p_n}{q_n}$ be the convergents to a real x and define $D_n = q_n x - p_n$ for all $n \geq 0$.*
Now let $n \geq 1$ be fixed. Then for all pairs of integers p and q with $0 < q \leq q_n$ we have

$$|qx - p| = |u||D_n| + |v||D_{n-1}|\,,$$

where

$$u = (-1)^{n+1} \det \begin{pmatrix} p & p_{n-1} \\ q & q_{n-1} \end{pmatrix} \quad \textit{and} \quad v = (-1)^{n+1} \det \begin{pmatrix} p_n & p \\ q_n & q \end{pmatrix}$$

uniquely solve the following system of linear equations:

$$\begin{cases} p = up_n + vp_{n-1} \\ q = uq_n + vq_{n-1} \end{cases}$$

Furthermore: u and v are integers, $uv \leq 0$ and if $q < q_n$, then $v \neq 0$.

A proof of Lemma 1.7 can be found in [1]. However, we will now prove Theorem 1.6.

Proof (Theorem 1.6). We first show that the denominator of any convergent is a best approximation. If $q_n = 1$, then q_n is trivially a best approximation, so we can assume that $q_n > 1$ and consider an integer q with $0 < q < q_n$. Let $p \in \mathbb{Z}$ such that $\|qx\| = |qx - p|$. By the Lemma, there are integers u and v with $v \neq 0$ such that

$$\begin{aligned} \|qx\| &= |qx - p| \\ &= |u||D_n| + |v||D_{n-1}| \\ &\geq |D_{n-1}| \\ &= |q_{n-1}x - p_{n-1}| \\ &= \|q_{n-1}x\|\,. \end{aligned} \tag{1}$$

The last inequality follows from the following observation:

Recall that

$$\left|x - \frac{p_{n-1}}{q_{n-1}}\right| < \frac{1}{q_n q_{n-1}},$$

and therefore $|q_{n-1}x - p_{n-1}| < \frac{1}{q_n} \leq \frac{1}{2}$, so p_{n-1} is the closest integer to $q_{n-1}x$. Similarly, $|q_n x - p_n| = \|q_n x\|$.

Before we continue, let us recall some other facts:

(i) $q_{n+1} = a_{n+1}q_n + q_{n-1} \geq q_n + q_{n-1}$.

(ii) $\frac{1}{q_n(q_n+q_{n+1})} < \left|x - \frac{p_n}{q_n}\right| \leq \frac{1}{q_n q_{n+1}} < \frac{1}{q_{n+1}^2}$.

Combining the results above with these two facts yields

$$\|q_n x\| = |q_n x - p_n| \leq \frac{1}{q_{n+1}} \leq \frac{1}{q_n + q_{n-1}} < |q_{n-1}x - p_{n-1}| \leq \|qx\|.$$

Therefore, the denominator of any convergent is a best approximation.

For the converse direction, assume that an integer $q > 0$ is a best approximation with $q_{n-1} < q < q_n$ for some n. By (1), we get $\|qx\| \geq \|q_{n-1}x\|$. However, by the definition of best approximation, we also know that $\|qx\| < \|q_{n-1}x\|$. We have reached a contradiction and the Theorem is proved. $\square$

Another Theorem that illustrates how well irrationals can be approximated by convergents, is the following due to Legendre.

Theorem 1.8 (Legendre)**.** *Let $x \in \mathbb{R} \setminus \mathbb{Q}$, and let p and q be corpime integers $(q > 0)$ with*

$$\left|x - \frac{p}{q}\right| \leq \frac{1}{2q^2}.$$

Then q is a best approximation to x and therefore the denominator of a convergent.

Proof. Take any integer q' such that $0 < q' < q$, and pick $p' \in \mathbb{Z}$ such that $|q'x - p'| = \|q'x\|$. Since p and q are coprime, we know that $\frac{p'}{q'} \neq \frac{p}{q}$. Therefore

$$\left|\frac{p'}{q'} - \frac{p}{q}\right| = \left|\frac{p'q - pq'}{qq'}\right| \geq \frac{1}{qq'},$$

and hence, using the triangle inequality:

$$\left|x - \frac{p'}{q'}\right| \geq \left|\frac{p'}{q'} - \frac{p}{q}\right| - \left|x - \frac{p}{q}\right| \geq \frac{1}{qq'} - \frac{1}{2q^2} > \frac{1}{2qq'}.$$

We conclude that

$$\left\|q'x\right\| = \left|q'x - p'\right| > \frac{1}{2q} \geq |qx - p| \geq \|qx\| \, .$$

$\square$

2 Hurwitz's Theorem

In this section, we treat a famous Theorem by Hurwitz that gives a bound to how well irrationals can be approximated (by convergents). In the proof of Hurwitz' Theorem, we will also see exactly for which number this bound is sharp.

Theorem 2.1 (Hurwitz, 1891)**.**

(i) For all $x \in \mathbb{R} \setminus \mathbb{Q}$ there are infinitely many pairs of integers p and q with $q > 0$ such that

$$\left|x - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}. \tag{2}$$

(ii) Furthermore, the constant $\frac{1}{\sqrt{5}}$ is optimal. This means, that for $(\sqrt{5}+\epsilon)^{-1}$, there are irrationals $x \in \mathbb{R} \setminus \mathbb{Q}$ such that the inequality (2) *only holds for finitely many pairs of integers.*

Proof. (i) We show that at least one out of any three consecutive convergents satisfies (2). Let $n \geq 2$ and assume the contrary, i.e. assume that

$$\left|x - \frac{p_{n-j}}{q_{n-j}}\right| \geq \frac{1}{\sqrt{5}q_{n-j}^2}, \qquad j = 0, 1, 2.$$

Using the fact that $|p_n q_{n-1} - q_n p_{n-1}| = \left|(-1)^{n+1}\right| = 1$, we get

$$\begin{aligned} \frac{1}{q_n q_{n-1}} = \frac{|p_n q_{n-1} - q_n p_{n-1}|}{|q_n q_{n-1}|} &= \left|\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}}\right| \\ &= \left|\frac{p_n}{q_n} - x\right| + \left|x - \frac{p_{n-1}}{q_{n-1}}\right| \\ &\geq \frac{q_{n-1}^2 + q_n^2}{\sqrt{5}q_{n-1}^2 q_n^2}. \end{aligned}$$

Multiplying this inequality by $(\sqrt{5}q_n^2)$ yields

$$\sqrt{5}\frac{q_n}{q_{n-1}} \geq 1 + \frac{q_n^2}{q_{n-1}^2}.$$

By rearranging the terms and introducing the variable $\lambda_n = \frac{q_n}{q_{n-1}}$, we can rewrite the above inequality to get

$$\lambda_n^2 - \sqrt{5}\lambda_n + 1 \leq 0.$$

Hence we find that $\lambda_n \leq \frac{1+\sqrt{5}}{2}$ (golden ratio), and since $\lambda_n = \frac{q_n}{q_{n-1}}$ is a rational number, this inequality is strict, so $\lambda_n < \frac{1+\sqrt{5}}{2}$. Analogously, we also get $\lambda_{n-1} < \frac{1+\sqrt{5}}{2}$.
By fact (i) recalled in the proof of Theorem 1.6, we get $q_n \geq q_{n-1} + q_{n-2}$, and hence $\lambda_n \geq 1 + \lambda_{n-1}^{-1}$. Therefore

$$\frac{1+\sqrt{5}}{2} > \lambda_n \geq 1 + \lambda_{n-1}^{-1} > 1 + \frac{1}{\frac{1+\sqrt{5}}{2}} = \frac{1+\sqrt{5}}{2},$$

and a contradiction has been reached.

(ii) We show that if we replace $\frac{1}{\sqrt{5}}$ by $\frac{1}{\sqrt{5}+\epsilon}$ for some $\epsilon > 0$, there are some irrational numbers for which the inequality (2) fails to be true for infinitely many pairs of integers p and q.
In fact, we prove that it fails for the golden ratio $\alpha_1 \coloneqq \frac{1+\sqrt{5}}{2}$.
Note that the golden ratio is a root of the polynomial

$$f(t) = t^2 - t - 1 = (t - \alpha_1)(t - \alpha_2),$$

where $\alpha_2 = \frac{1-\sqrt{5}}{2}$.

By assumption, there is a sequence of rational numbers $(\frac{p_k}{q_k})_{k\geq 1}$, such that

$$\forall k \geq 1 : \left|\alpha_1 - \frac{p_k}{q_k}\right| < \frac{1}{(\sqrt{5}+\epsilon)q_k^2}.$$

Therefore we get:

$$\begin{aligned}\left|\alpha_2 - \frac{p_k}{q_k}\right| &\leq |\alpha_2 - \alpha_1| + \left|\alpha_1 - \frac{p_k}{q_k}\right| \\ &< \sqrt{5} + \frac{1}{(\sqrt{5}+\epsilon)q_k^2} \\ &< \sqrt{5} + \epsilon, \quad \text{for all } k \text{ sufficiently large.}\end{aligned}$$

Consequently,

$$\begin{aligned}
\left|f(\frac{p_k}{q_k})\right| &= \left|\frac{p_k}{q_k} - \alpha_1\right|\left|\frac{p_k}{q_k} - \alpha_2\right| \\
&< \frac{\sqrt{5}+\epsilon}{(\sqrt{5}+\epsilon)q_k^2}, \quad \text{for all } k \text{ sufficiently large} \\
&= \frac{1}{q_k^2}.
\end{aligned}$$

Therefore, for all sufficiently large k, we have $\left|q_k^2 f(\frac{p_k}{q_k})\right| < 1$. We also know that $\left|q_k^2 f(\frac{p_k}{q_k})\right| > 0$, since $q_k \neq 0$ and both roots of f are irrational. However,

$$\left|q_k^2 f(\frac{p_k}{q_k})\right| = \left|q_k^2(\frac{p_k^2}{q_k^2} - \frac{p_k}{q_k} - 1)\right| = \left|p_k^2 - p_k q_k - q_k^2\right| \in \mathbb{Z},$$

contradicting $0 < \left|q_k^2 f(\frac{p_k}{q_k})\right| < 1$.

□

3 Lagrange spectrum

Definition 3.1. For a real number x, define $\nu(x) = \liminf_{q\to\infty} q\|qx\|$.

Remark 3.2. By Hurwitz' Theorem 2.1, we know that $\nu(x) \in [0, \frac{1}{\sqrt{5}}]$ and $\nu(\frac{1+\sqrt{5}}{2}) = \frac{1}{\sqrt{5}}$. Furthermore, it can be shown that $\nu(x) = \nu(y)$ for any two equivalent numbers x and y.

Theorem 3.3 (Markov)**.** *There exists a sequence*

$$\mu_1 = \frac{1}{\sqrt{5}} > \mu_2 = \frac{1}{2\sqrt{2}} > \mu_3 = \frac{5}{\sqrt{221}} > \mu_4 > \dots$$

with $\lim_{n\to\infty} \mu_n = \frac{1}{3}$*, such that for all* $\mu > \frac{1}{3}$ *there exists* $x \in \mathbb{R}$ *with* $\nu(x) = \mu$ *if and only if* $\mu = \mu_k$ *for some* $k \geq 1$.
Furthermore, for each $k \geq 1$ *there is only a finite number of equivalence classes of* x *with* $\nu(x) = \mu_k$*, and there are uncountably many* $x \in \mathbb{R}$ *with* $\nu(x) = \frac{1}{3}$.

Definition 3.4. The set $\mathcal{L} = \{\nu(x) : x \in \mathbb{R}\} \subset [0, \frac{1}{\sqrt{5}}]$ is called the *Lagrange spectrum*.

By Theorem 3.3, we see that $\mathcal{L} \cap (\frac{1}{3}, \frac{1}{\sqrt{5}}] = \{\mu_1, \mu_2, \dots\}$ is discrete. However, $\mathcal{L} \setminus (\frac{1}{3}, \frac{1}{\sqrt{5}}]$ is not discrete; Marshall Hall proved in 1947 that there exists $c > 0$ such that $[0, c] \subset \mathcal{L}$.

In fact, in 1975 Freiman determined $\max\{c > 0 : [0, c] \subset \mathcal{L}\}$. It is given by

$$F = \frac{491993569}{2221564096 + 283748\sqrt{462}} = 0.220856\dots$$

Definition 3.5. $F^{-1} = 4.527829\dots$ is called *Freiman's constant*.

In conclusion, $\mathcal{L} \cap (\frac{1}{3}, \frac{1}{\sqrt{5}}] = \{\mu_1, \mu_2, \dots\}$ (discrete) and $\mathcal{L} \cap [0, F] = [0, F]$ (continuous) are well-understood. However, the region $\mathcal{L} \cap (F, \frac{1}{3}]$ is a little chaotic. For example, Hightower proved in 1970 that there are countably many disjoint intervals $I_k \subset (F, \frac{1}{3}]$, such that $\mathcal{L} \cap I_k = \emptyset$ for all k, but there exists a point $x \in \mathcal{L}$ between any two intervals I_{k_1} and I_{k_2}. Furthermore, it is known that $\mathcal{L} \cap (F, \frac{1}{3}]$ has a fractal structure.

References

[1] Victor Beresnevich. *Metric Number Theory*. Lecture Notes. 2014.

[2] G.H. Hardy and E.M. Wright. *An Introduction to the Theory of Numbers*.

[3] A. Ya. Khinchin. *Continued Fractions*. Dover Publications.

[4] Carlos Gustavo Moreira. *Geometric properties of the Markov and Lagrange spectra*. 2016. arXiv: `1612.05782 [math.NT]`.