Hermite-Padé Approximants and Transcendence of e, π

Maxim L. Yattselev

IUPUI
SCHOOL OF SCIENCE
Department of Mathematical Sciences

Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo, São Carlos
May 26th, 2022

Algebraic and Transcendental Numbers

A number α is called *algebraic (of degree d)* if there exists a polynomial

$$f(x) = n_d x^d + n_{d-1} x^{d-1} + \cdots + n_0$$

with *integer* coefficients, $n_d \neq 0$, such that $f(\alpha) = 0$ (and there are no such polynomial of smaller degree).

A number α is *transcendental* if it is not algebraic.

Theorem (Liouville 1844)

If a non-rational number α is algebraic of degree d, then

$$\frac{C}{q^d} \leq \left|\alpha - \frac{p}{q}\right|$$

for some constant C and any rational number p/q, $q \geq 1$.
That is, α *cannot be approximated well* by rational numbers.

Let $f(x)$ be the corresponding minimal polynomial with integer coefficients and A be such that $f(x) \neq 0$ for $x \in (\alpha - A, \alpha + A) \setminus \{\alpha\}$. Then

$$\left|\alpha - \frac{p}{q}\right| \geq A \quad \Rightarrow \quad \left|\alpha - \frac{p}{q}\right| \geq \frac{A}{q^d}.$$

If $|\alpha - p/q| < A$, then

$$\frac{1}{q^d} \leq |f(p/q)| = |f(p/q) - f(\alpha)| \leq \max_{x \in [\alpha - A, \alpha + A]} |f'(x)| \left|\alpha - \frac{p}{q}\right|.$$

Theorem (Liouville 1844)

$\alpha = \sum_{n\geq 1} \frac{1}{10^{n!}}$ is transcendental.

Let

$$\frac{p}{q} = \sum_{n=1}^{N} \frac{1}{10^{n!}} = \frac{p}{10^{N!}}.$$

Set $r = \dfrac{1}{10^{(N+1)!}}$. Then

$$\begin{aligned} \alpha - \frac{p}{q} &= \sum_{n>N} \frac{1}{10^{n!}} = r\left(1 + \frac{1}{r^{(N+2)-1}} + \frac{1}{r^{(N+2)(N+3)-1}} + \cdots\right) \\ &< r\left(1 + \frac{1}{r} + \frac{1}{r^2} + \cdots\right) = \frac{r}{1-r} < 2r = \frac{2}{q^{N+1}}. \end{aligned}$$

Irrationality: Criterion

Lemma

α is irrational if for any $\varepsilon > 0$ there exist a pair of integers (q, p) such that $0 < \left|q\alpha - p\right| \leq \varepsilon$.

If $\alpha = u/v$, then

$$\varepsilon \geq |q\alpha - p| = \frac{|qu - pv|}{v} \geq \frac{1}{v}.$$

Irrationality: Criterion

Lemma

α is irrational if for any $\varepsilon > 0$ there exist a pair of integers (q, p) such that $0 < \left|q\alpha - p\right| \leq \varepsilon$.

If $\alpha = u/v$, then

$$\varepsilon \geq |q\alpha - p| = \frac{|qu - pv|}{v} \geq \frac{1}{v}.$$

To prove that e is irrational, we need to find a sequence of rational numbers that approximate e and are never equal to e.

Hermite realized that he can construct such a sequence by cleverly approximating e^x by rational functions.

Padé Approximants: Definition

Let

$$F(x) = f_0 + f_1 x + f_2 x^2 + \cdots + f_k x^k + \cdots$$

be a convergent power series. Let

$$Q(x) = q_0 + q_1 x + \cdots + q_n x^n$$

and

$$P(x) = p_0 + p_1 x + \cdots + p_m x^m$$

be polynomials such that

$$Q(x)F(x) - P(x) = \mathcal{O}\left(x^{m+n+1}\right)$$

A rational function $P(x)/Q(x)$ is called a Padé approximant of type (m, n) to $F(x)$ (Padé was a student of Hermite).

Padé Approximants: Existence

It holds that

$$\begin{aligned} Q(x)F(x) &= f_0q_0 + (f_0q_1 + f_1q_0)x + (f_0q_2 + f_1q_1 + f_2q_0)x^2 + \cdots \\ &= \sum_{k=0}^{\infty}\left(\sum_{i=1}^{\min\{k,n\}} f_{k-i}q_k\right)x^k. \end{aligned}$$

Therefore,

$$p_k := \sum_{i=1}^{\min\{k,n\}} f_{k-i}q_k \quad \text{and} \quad \sum_{i=1}^{\min\{k,n\}} f_{k-i}q_i = 0$$

for k from 0 to m and and for k from $m+1$ to $m+n+1$, respectively.

The latter is a linear system of n equations with $n+1$ unknowns. Such a system always has a solution.

Padé Approximants: Uniqueness

A solution may not be unique, but the ratio $P(x)/Q(x)$ *always is*.

Indeed, let $x^{k_1}Q_1(x), x^{k_1}P_1(x)$ and $x^{k_2}Q_2(x), x^{k_2}P_2(x)$ be solutions. Then

$$F(x) - \frac{P_j(x)}{Q_j(x)} = \mathcal{O}\left(x^{m+n+1-k_j}\right).$$

Therefore,

$$\frac{P_1(x)}{Q_1(x)} - \frac{P_2(x)}{Q_2(x)} = \mathcal{O}\left(x^{m+n+1-\max k_j}\right).$$

However,

$$\deg\left(P_1Q_2 - P_2Q_1\right) \leq m + n - k_1 - k_2 < m + n + 1 - \max k_j.$$

Differentiation Operator

For any convergent power series $f(x)$, let

$$\mathcal{I}[f(x)] = f(x) \quad \text{and} \quad \mathcal{D}[f(x)] = f'(x).$$

Observe that

$$\mathcal{D}[e^x f(x)] = e^x f(x) + e^x f'(x) = e^x (\mathcal{I} + \mathcal{D})[f(x)],$$

while

$$\begin{aligned} \mathcal{D}^2[e^x f(x)] &= e^x f(x) + 2e^x f'(x) + e^x f''(x) \\ &= e^x \left(\mathcal{I} + 2\mathcal{D} + \mathcal{D}^2\right)[f(x)] = e^x (\mathcal{I} + \mathcal{D})^2 [f(x)] \end{aligned}$$

and so on. Notice also that

$$\mathcal{D}\left[\mathcal{O}\left(x^{k+1}\right)\right] = \mathcal{O}\left(x^k\right).$$

Padé Approximants to e^x

Let $P(x)/Q(x)$ be the Padé approximant of type (m, n) to e^x. Then

$$Q(x)e^x - P(x) = \mathcal{O}\left(x^{m+n+1}\right),$$

and therefore it holds that

$$\mathcal{D}^{m+1}\left[Q(x)e^x - P(x)\right] = e^x(\mathcal{I} + \mathcal{D})^{m+1}[Q(x)] = \mathcal{O}\left(x^n\right).$$

In particular,

$$(\mathcal{I} + \mathcal{D})^{m+1}[Q(x)] = \kappa x^n$$

where κ is the leading coefficient $Q(x)$. Hence,

$$Q(x) = (\mathcal{I} + \mathcal{D})^{-(m+1)}\left[\kappa x^n\right] = \sum_{k=0}^{\infty}(-1)^k\binom{m+k}{k}\mathcal{D}^k\left[\kappa x^n\right].$$

Padé Approximants to e^x

Normalizing $Q(0) = 1$, we get that

$$Q(x) = \sum_{k=0}^{n} \binom{n}{k} \frac{(m+n-k)!}{(m+n)!} (-x)^k$$

(in this case $\kappa = (-1)^n m!/(n+m)!$) and

$$P(x) = \sum_{k=0}^{m} \binom{m}{k} \frac{(m+n-k)!}{(m+n)!} x^k$$

Observe also that

$$\frac{(m+n)!}{m!} Q(\ell) \quad \text{and} \quad \frac{(m+n)!}{n!} P(\ell)$$

are integers for $\ell \in \mathbb{Z}$.

Recall that

$$f(x) - T_m[f(x)] = \frac{1}{m!}\int_0^x (x-t)^m f^{(m+1)}(t)dt,$$

where $T_m[f(x)]$ is the m-th McLaurin polynomial of $f(x)$. Notice that

$$P(x) = T_m\left[Q(x)e^x\right].$$

Recall also that

$$\kappa x^n e^x = \mathcal{D}^{m+1}\left[Q(x)e^x - P(x)\right] = \mathcal{D}^{m+1}\left[Q(x)e^x\right].$$

Since $\kappa = (-1)^n m!/(n+m)!$,

$$\begin{aligned} Q(x)e^x - P(x) &= \frac{\kappa}{m!}\int_0^x (x-t)^m t^n e^t\,dt \\ &= \frac{\kappa}{m!}\int_0^1 (x-sx)^m (sx)^n e^{sx} x\,ds \\ &= \frac{(-1)^n}{(n+m)!} x^{m+n+1}\int_0^1 (1-s)^m s^n e^{sx}\,ds. \end{aligned}$$

Irrationality of e^ℓ

Theorem

e^ℓ is irrational for any integer ℓ.

Let $P(x)/Q(x)$ be the Padé approximant of type (n, n) to e^x. Set

$$p := \frac{(2n)!}{n!}P(\ell) \quad \text{and} \quad q := \frac{(2n)!}{n!}Q(\ell).$$

Then p, q are integers and

$$qe^\ell - p = \frac{(-1)^n}{n!}\ell^{2n+1}\int_0^1 (1-s)^n s^n e^{s\ell}\, ds.$$

Hence,

$$0 < \left|qe^\ell - p\right| < \frac{\ell^{2n+1}e^\ell}{n!}.$$

Transcendance: Criterion

Lemma

α is transcendental if for any $d \in \mathbb{N}$ and $\varepsilon > 0$ there exist $d+1$ linearly independent vectors of integers $(q_j, p_{j1}, \dots, p_{jd})$, $j = \overline{0, d}$, such that

$$\left| q_j \alpha^k - p_{jk} \right| \leq \varepsilon \quad \text{for each } \ j \ \text{ and } \ k.$$

If α is algebraic of degree d, then $n_0 + n_1\alpha + \cdots + n_d\alpha^d = 0$, $n_k \in \mathbb{Z}$. Hence,

$$0 = \sum_{k=1}^{d} n_k (q_j \alpha^k - p_{jk}) + n_0 q_j + \sum_{k=1}^{d} n_k p_{jk}.$$

By linear independence, there exists i such that

$$n_0 q_i + n_1 p_{i1} + \cdots + n_d p_{id} \neq 0.$$

Therefore,

$$1 \leq \left| \sum_{k=1}^{d} n_k (q_i \alpha^k - p_{ik}) \right| \leq \varepsilon \sum_{k=1}^{d} |n_k|.$$

Hermite-Padé Approximants

Let $F_1(x), \ldots, F_d(x)$ be convergent power series at the origin.

Let $n_0, n_1, \ldots, n_d$ be non-negative integers.

Set $N := n_0 + \cdots + n_d$ and consider the following system of equations

$$R_k(x) := Q(x)F_k(x) - P_k(x) = \mathcal{O}\left(x^{N+1}\right)$$

where $\deg(Q) \leq N - n_0$ and $\deg(P_k) \leq N - n_k$.

Such polynomials exist (their coefficients are obtained from a linear system which is always solvable), but are not necessarily unique.

The d-tuple of rational functions $P_1/Q, \ldots, P_d/Q$ is called an *Hermite-Padé approximant of type II*.

Hermite Padé Approximants for $e^x, \dots, e^{dx}$

Proposition

The d-tuple of Hermite-Padé approximants for the system $e^x, \dots, e^{dx}$ is *unique* and is given up to the normalization by the formulae

$$
\begin{aligned}
Q(x) &= f^{(N)}(0) + f^{(N-1)}(0)x + \cdots + f^{(n_0)}(0)x^{N-n_0} \\
P_k(x) &= f^{(N)}(k) + f^{(N-1)}(k)x + \cdots + f^{(n_k)}(k)x^{N-n_k} \\
R_k(x) &= x^{N+1} \int_0^k f(s)e^{x(k-s)}\,ds
\end{aligned}
$$

where $f(s) = s^{n_0}(s-1)^{n_1} \cdots (s-d)^{n_d}$.

Moreover, the polynomials $(1/n_0!)Q(x)$ and $(1/n_k!)P_k(x)$ have integer coefficients.

Hermite-Padé Approximants for $e^x, \ldots, e^{dx}$: Linear Independence

Proposition

$$\Delta(x) := \begin{vmatrix} Q_0(x) & P_{01}(x) & \cdots & P_{0d}(x) \\ \vdots & \vdots & \ddots & \vdots \\ Q_d(x) & P_{d1}(x) & \cdots & P_{dd}(x) \end{vmatrix} = \kappa x^{dN},$$

$\kappa \neq 0$, where Hermite-Padé approximants are associated to the following system of indices:

$$\begin{vmatrix} n_0 - 1 & n_1 & \cdots & n_d \\ n_0 & n_1 - 1 & \cdots & n_d \\ \vdots & \vdots & \ddots & \vdots \\ n_0 & n_1 & \cdots & n_d - 1 \end{vmatrix}.$$

Hermite-Padé Approximants for $e^x, \ldots, e^{dx}$: Linear Independence

Let $N = n_0 + n_1 + \cdots + n_d$. Since the system of indices is

$$\begin{vmatrix} n_0 - 1 & n_1 & \cdots & n_d \\ n_0 & n_1 - 1 & \cdots & n_d \\ \vdots & \vdots & \ddots & \vdots \\ n_0 & n_1 & \cdots & n_d - 1 \end{vmatrix},$$

it holds that

$$\deg(\Delta) = \begin{vmatrix} N - n_0 & N - n_1 - 1 & \cdots & N - n_d - 1 \\ N - n_0 - 1 & N - n_1 & \cdots & N - n_d - 1 \\ \vdots & \vdots & \ddots & \vdots \\ N - n_0 - 1 & N - n_1 - 1 & \cdots & N - n_d \end{vmatrix} = dN$$

and that the coefficient next to x^{dN} is the product of leading coefficients of the elements of the main diagonal and therefore is not zero.

On the other hand,

$$\Delta(x) = \begin{vmatrix} Q_0(x) & P_{01}(x) - e^x Q_0(x) & \cdots & P_{0d}(x) - e^{dx} Q_0(x) \\ \vdots & \vdots & \ddots & \vdots \\ Q_d(x) & P_{d1}(x) - e^x Q_d(x) & \cdots & P_{dd}(x) - e^{dx} Q_d(x) \end{vmatrix}$$

$$= \begin{vmatrix} Q_0(x) & \mathcal{O}\left(x^N\right) & \cdots & \mathcal{O}\left(x^N\right) \\ \vdots & \vdots & \ddots & \vdots \\ Q_d(x) & \mathcal{O}\left(x^N\right) & \cdots & \mathcal{O}\left(x^N\right) \end{vmatrix} = \mathcal{O}\left(x^{dN}\right).$$

Theorem (Hermite 1873)

e is transcendental.

For any $d, n \in \mathbb{N}$, consider $(d+1)$-tuples of the Hermite-Padé approximants

$$\Big(Q_j(x), P_{j1}(x), \ldots, P_{jd}(x)\Big)$$

for $e^x, \ldots, e^{dx}$ associated with the matrix of indices

$$\begin{vmatrix} n-1 & n & \cdots & n \\ n & n-1 & \cdots & n \\ \vdots & \vdots & \ddots & \vdots \\ n & n & \cdots & n-1 \end{vmatrix}.$$

Set

$$q_j := \frac{Q_j(1)}{(n-1)!} \quad \text{and} \quad p_{jk} := \frac{P_{jk}(1)}{(n-1)!}.$$

Then q_j, p_{jk} are integers,

$$\left|q_j e^k - p_{jk}\right| = \frac{1}{(n-1)!}\left|\int_0^k \frac{f(s)}{s-j} e^{k-s}\, ds\right| \leq \frac{e^d d^{dn+1}}{(n-1)!},$$

where $f(s) = s^n (s-1)^n \cdots (s-d)^n$, and

$$\begin{vmatrix} q_0 & p_{01} & \cdots & p_{0d} \\ \vdots & \vdots & \ddots & \vdots \\ q_d & p_{d1} & \cdots & p_{dd} \end{vmatrix} = \frac{\Delta(1)}{((n-1)!)^{d+1}} \neq 0.$$

Theorem (Lindemann 1882)

π is transcendental.

Assume π is algebraic, in which case $\mathrm{i}\pi$ is algebraic as well.

Let g be the polynomial with Gaussian integer coefficients such that $g(\mathrm{i}\pi) = 0$.

Denote by $\alpha_1, \ldots, \alpha_D$ the zeros of g. Then

$$0 = (e^{\alpha_1} + 1) \cdots (e^{\alpha_D} + 1) = e^{\beta_1} + \cdots + e^{\beta_d} + 2^D - d,$$

where β_l are combinations $\epsilon_1 \alpha_1 + \cdots + \epsilon_D \alpha_D$ with $\epsilon_l \in \{0, 1\}$ that are not equal to zero.

Transcendence of π

For any $n \in \mathbb{N}$, consider $(d+1)$-tuples of the Hermite-Padé approximants

$$\Big(Q_j(x), P_{j1}(x), \ldots, P_{jd}(x)\Big)$$

for $e^{\beta_1 x}, \ldots, e^{\beta_d x}$ associated with the matrix of indices

$$\begin{vmatrix} n-1 & n & \cdots & n \\ n & n-1 & \cdots & n \\ \vdots & \vdots & \ddots & \vdots \\ n & n & \cdots & n-1 \end{vmatrix}.$$

Let $N = (d+1)n - 1$ and $n_{jk} = n - \delta_{jk}$. Then

$$\begin{aligned} Q_j(x) &= f_j^{(N)}(0) + f_j^{(N-1)}(0)x + \cdots + f_j^{(n_{j0})}(0)x^{N-n_{j0}} \\ P_{jk}(x) &= f_j^{(N)}(\beta_k) + f_j^{(N-1)}(\beta_k)x + \cdots + f_j^{(n_{jk})}(\beta_k)x^{N-n_{jk}} \\ R_{jk}(x) &= x^{N+1}\int_0^{\beta_k} f_j(s)e^{x(\beta_k - s)}\,ds \end{aligned}$$

where $f_j(s) = s^{n_{j0}}(s-\beta_1)^{n_{j1}} \cdots (s-\beta_d)^{n_{jd}}$.

Set

$$q_j := \frac{\varkappa Q_j(1)}{(n-1)!} \quad \text{and} \quad p_{jk} := \frac{\varkappa P_{jk}(1)}{(n-1)!},$$

where $\varkappa$ is the leading coefficient of g.

Similarly to the previous proof, one can show that vectors $(q_j \ p_{1j} \ \dots \ p_{dj})$ are linearly independent. Hence, there exists an index $i = i(n)$ such that

$$q_i(2^D - d) + \sum_{k=1}^{d} p_{ik} \neq 0.$$

Using the fact that the symmetric functions of $\alpha_1, \dots, \alpha_d$ are integers, one can also argue that the sum above is an integer. Then

$$0 = \sum_{k=1}^{d} q_i e^{\beta_k} + q_i(2^D - d) = \sum_{k=1}^{d} \left(q_i e^{\beta_k} - p_{ik}\right) + q_i(2^D - d) + \sum_{k=1}^{d} p_{ik},$$

where the first sum can be made arbitrarily small by increasing n and the remaining summands always add up to a non-zero integer.