The hypergeometric method

Michael Bennett

University of British Columbia

BIRS Summer School : June 2012

One way to show that a number is irrational

Suppose that we are given a real number θ that we wish to prove to be irrational. One way to do this is to find a sequence of distinct rational approximations p_n/q_n to θ (here, p_n and q_n are integers) with the property that there exist positive real numbers α, β, a and b with $\alpha, \beta > 1$,

$$|q_n| < a \cdot \alpha^n, \text{ and } |q_n\theta - p_n| < b \cdot \beta^{-n},$$

for all positive integers n.

What does this do?

If we can construct such a sequence, we in fact get rather more. Namely, we obtain the inequality

$$\left| \theta - \frac{p}{q} \right| > \left(2\, a\, \alpha\, (2\, b\, \beta)^{\lambda} \right)^{-1} |q|^{-1-\lambda} \quad \text{for} \quad \lambda = \frac{\log \alpha}{\log \beta}, \quad (*)$$

valid for *all* integers p and $q \neq 0$ (at least provided $|q| > 1/2b$). To see this, note that

$$\left| \theta - \frac{p}{q} \right| \geq \left| \frac{p_n}{q_n} - \frac{p}{q} \right| - \left| \theta - \frac{p_n}{q_n} \right|,$$

and hence if $p/q \neq p_n/q_n$, we have

$$\left| \theta - \frac{p}{q} \right| > \frac{1}{|q_n|} \left(\frac{1}{|q|} - \frac{b}{\beta^n} \right).$$

A lower bound

Now we choose n minimal such that $\beta^n \geq 2\,b\,|q|$ (since we assume $|q| > 1/2b$, n is a positive integer). Then

$$\beta^{n-1} < 2\,b\,|q| \leq \beta^n$$

and so

$$\left|\theta - \frac{p}{q}\right| > \frac{1}{2\,|q\,q_n|} > \frac{1}{2\,|q|\,a\,\alpha^n} = \frac{1}{2\,|q|\,a\,\beta^{\lambda n}} > \frac{1}{2\,|q|\,a\,(2|q|\,b\,\beta)^{\lambda}}.$$

If instead we have $p/q = p_n/q_n$ for our desired choice of n, we argue similarly, only with n replaced by $n+1$ (whereby the fact that our approximations are distinct guarantees that $p/q \neq p_{n+1}/q_{n+1}$). The slightly weaker constant in (*) results from this case.

In summary

If

$$|q_n| < a \cdot \alpha^n, \quad \text{and} \quad |q_n\theta - p_n| < b \cdot \beta^{-n},$$

for all positive integers n, then

$$\left|\theta - \frac{p}{q}\right| > \left(2\,a\,\alpha\,(2\,b\,\beta)^{\lambda}\right)^{-1} |q|^{-1-\lambda} \quad \text{for} \quad \lambda = \frac{\log \alpha}{\log \beta}, \quad (*)$$

for all integers p and $q \neq 0$, with $|q| > 1/2b$.

Irrationality measures

An inequality of the shape

$$\left|\theta - \frac{p}{q}\right| > |q|^{-\kappa},$$

valid for suitably large integers p and q is termed an *irrationality measure*. For real transcendental θ, any such measure is in some sense nontrivial. For algebraic θ, say of degree n, however, Liouville's theorem provides a "trivial" lower bound of n for κ.

The hypergeometric method

In this talk, we will provide an oversimplified account of one technique for generating, for certain irrational θ, sequences p_n/q_n with the properties described above. This approach is sometimes called *hypergeometric method* and dates back to work of Thue.

Irrationality of π

The following argument is due to Beukers, inspired by Apery's proof of the irrationality of $\zeta(3)$.

We will sketch a proof of the irrationality of $\zeta(2) = \pi^2/6$. This is by no means the original proof, but it is rather instructive.

To begin, we note the identity

$$\int_0^1 \int_0^1 \frac{x^r y^s}{1-xy} dx\, dy = \sum_{k=0}^{\infty} \frac{1}{(k+r+1)(k+s+1)}.$$

To see this, just expand $1/(1-xy)$ as a geometric series, substitute and carry out the integrations.

A dichotomy

Assuming $r = s$, we have

$$\int_0^1 \int_0^1 \frac{(xy)^r}{1-xy} = \zeta(2) - \left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{r^2}\right).$$

If, instead, we consider a similar integral, with distinct powers of x and y in the numerator of the integrand, we find, via partial fractions and telescoping,

$$\int_0^1 \int_0^1 \frac{x^r y^s}{1-xy} dx\, dy = \frac{1}{r-s}\left(\frac{1}{s+1} + \cdots + \frac{1}{r}\right).$$

From rationals to integers

Defining

$$L(r) = \text{lcm}(1, 2, \ldots, r),$$

it follows that

$$L(r)^2 \int_0^1 \int_0^1 \frac{x^r y^s}{1-xy} dx\, dy$$

is thus, for each $r > s$, a positive integer.

Approximating $\zeta(2)$

Combining these facts, we find that, for any polynomial $P(x,y)$ with integer coefficients,

$$\int_0^1 \int_0^1 \frac{P(x,y)}{1-xy} dx\, dy = A\zeta(2) - B,$$

for *rational* A and B. We will, through careful choice of a family of such $P(x,y)$, construct our sequences of integers p_n and q_n such that p_n/q_n is a suitably good approximation to $\zeta(2)$.

A choice for $P(x,y)$

Let us take

$$P(x,y) = (1-y)^n P_n(x),$$

where

$$P_n(x) = \frac{1}{n!}\left(\frac{d}{dx}\right)^n \left(x^n(1-x)^n\right).$$

Then, as noted before, we have

$$L(n)^2 \int_0^1 \int_0^1 \frac{(1-y)^n P_n(x)}{1-xy} dx\, dy = q_n \zeta(2) - p_n, \qquad (1)$$

with p_n, q_n rational integers.

Why make this choice?

One of the main reasons for this choice is that the integrand of the left-hand-side of (1) is now extremely small, while the coefficients of the numerator of the integrand do not grow too quickly. In fact, an n-fold integration by parts shows us that

$$\int_0^1 \int_0^1 \frac{(1-y)^n P_n(x)}{1-xy} dx\, dy$$

is equal to

$$\pm \int_0^1 \int_0^1 \frac{y^n(1-y)^n x^n (1-x)^n}{(1-xy)^{n+1}} dx\, dy.$$

Why make this choice?

Since

$$\frac{y(1-y)x(1-x)}{1-xy} \leq \left(\frac{\sqrt{5}-1}{2}\right)^5, \quad \text{for } 0 \leq x, y \leq 1,$$

we thus have

$$0 < |q_n \zeta(2) - p_n| \leq L(n)^2 \left(\frac{\sqrt{5}-1}{2}\right)^{5n} \int_0^1 \int_0^1 \frac{dx dy}{1-xy}.$$

The finale

Since, $\log L(n) \sim n$ (via the Prime Number Theorem), and since

$$e^2 \left(\frac{\sqrt{5}-1}{2} \right)^5 < 2/3,$$

it follows that $\zeta(2)$ is irrational.

The general idea

If we work a little more carefully, we can estimate the growth of p_n and q_n and get an irrationality measure for $\zeta(2)$ (and hence for π) from this argument. The thing to take away from this "example" is the notion of constructing our approximating sequences p_n/q_n via specialization of rational functions.

Padé approximants

Given a formal power series $f(z)$ and positive integers r and s, it is an exercise in linear algebra to deduce, for fixed integer n, the existence of nonzero polynomials $P_{r,s}(z)$ and $Q_{r,s}(z)$ with rational integer coefficients and degrees r and s, respectively, such that

$$P_{r,s}(z) - f(z)\, Q_{r,s}(z) = z^{r+s+1} E_{r,s}(z)$$

where $E_{r,s}(z)$ is a power series in z (let's not worry too much about convergence!). In certain situations, these *Padé approximants* (which are unique up to scaling) can be written down in explicit fashion.

Diagonal Padé approximants

Given a formal power series $f(z)$ and positive integers r and s, it is an exercise in linear algebra to deduce, for fixed integer n, the existence of nonzero polynomials $P_n(z)$ and $Q_n(z)$ with rational integer coefficients and degree n, such that

$$P_n(z) - f(z)\, Q_n(z) = z^{2n+1} E_n(z)$$

where $E_n(z)$ is a power series in z (let's not worry too much about convergence!). In certain situations, these *Padé approximants* (which are unique up to scaling) can be written down in explicit fashion.

Padé approximants to the binomial function

Such is the case for $f(z) = (1-z)^{1/m}$. Indeed, if we define

$$P_n(z) = \sum_{k=0}^{n} \binom{n+1/m}{k} \binom{2n-k}{n} (-z)^k$$

and

$$Q_n(z) = \sum_{k=0}^{n} \binom{n-1/m}{k} \binom{2n-k}{n} (-z)^k,$$

then there exists a power series $E_n(z)$ such that for all complex z with $|z| < 1$,

$$P_n(z) - (1-z)^{1/m} Q_n(z) = z^{2n+1} E_n(z). \tag{2}$$

What's in a name?

A *hypergeometric function* is a power series of the shape

$$1+\sum_{n=1}^{\infty}\frac{\alpha(\alpha+1)\cdots(\alpha+n-1)\beta(\beta+1)\cdots(\beta+n-1)}{\gamma(\gamma+1)\cdots(\gamma+n-1)n!}z^n.$$

We'll call such function $F(\alpha,\beta,\gamma,z)$. Here z is a complex variable and α, β and γ are complex constants. If α or β is a non-positive integer and m is the smallest integer such that

$$\alpha(\alpha+1)\cdots(\alpha+m)\beta(\beta+1)\cdots(\beta+m)=0,$$

then $F(\alpha,\beta,\gamma,z)$ is a polynomial in z of degree m.

The hypergeometric differential equation

It is worth noting that the hypergeometric function $F(\alpha, \beta, \gamma, z)$ satisfies the differential equation

$$z(1-z)\frac{d^2F}{dz^2} + (\gamma - (1+\alpha+\beta)z)\frac{dF}{dz} - \alpha\beta F = 0. \qquad (3)$$

Padé approximants as hypergeometric functions

Note that, in terms of hypergeometric functions,

$$P_n(z) = \binom{2n}{n} F(-1/m - n, -n, -2n, z)$$

and

$$Q_n(z) = \binom{2n}{n} F(1/m - n, -n, -2n, z),$$

An application of differential equations

Another way to see (4), then, is to note that the functions

$$P_n(z), \ (1-z)^{1/m} Q_n(z)$$

and

$$z^{2n+1} F(n+1, n+(m-1)/m, 2n+2, z)$$

each satisfy (3) with $\alpha = -1/m - n$, $\beta = -n$, $\gamma = -2n$, and hence, it is not too difficult to show, are linearly dependent over the rationals. Finding the dependency is then a reasonably easy exercise.

Most of the functions for which we are able to explicitly determine families of Padé approximants are special cases of the hypergeometric function.

Padé approximants to the binomial function : $m = 3$

In the case of $f(z) = (1-z)^{1/3}$, setting

$$P_n(z) = \sum_{k=0}^{n} \binom{n+1/3}{k} \binom{2n-k}{n} (-z)^k$$

and

$$Q_n(z) = \sum_{k=0}^{n} \binom{n-1/3}{k} \binom{2n-k}{n} (-z)^k,$$

there thus exists a power series $E_n(z)$ such that for all complex z with $|z| < 1$,

$$P_n(z) - (1-z)^{1/3} Q_n(z) = z^{2n+1} E_n(z). \qquad (4)$$

Some amazing facts

For later use, it is worth noting the following results.

LEMMA 1 : Let n be a positive integer and suppose that z is a complex number with $|1-z| \leq 1$. Then
(i) We have

$$|P_n(z)| < 4^n, \; |Q_n(z)| < 4^n$$

and

$$|E_n(z)| < 4^{-n}(1-|z|)^{-\frac{1}{2}(2n+1)}.$$

(ii) For all complex numbers $z \neq 0$, we have

$$P_n(z)Q_{n+1}(z) \neq P_{n+1}(z)Q_n(z).$$

Some amazing facts (continued)

(iii) If k is a nonnegative integer, then

$$3^{k+[k/2]} \binom{n+1/3}{k}$$

is an integer.

(iv) If we define G_n to be the largest positive integer such that

$$\frac{3^{n+[n/2]}P_n(z)}{G_n} \quad \text{and} \quad \frac{3^{n+[n/2]}Q_n(z)}{G_n}$$

are both polynomials with integer coefficients, then

$$G_n > \frac{1}{42}\, 2^n,$$

for all positive integers n.

An example

Let's substitute $z = -1/5831$ into

$$P_n(z) - (1-z)^{1/3} Q_n(z) = z^{2n+1} E_n(z).$$

Then for

$$\theta = \left(1 + \frac{1}{5831}\right)^{1/3},$$

we have

$$P_n - \theta Q_n = I_n,$$

with $P_n, Q_n \in \mathbb{Z}$. Specifically ...

An example (continued)

Specifically,

$$P_n - \theta Q_n = I_n,$$

with

$$P_n = \frac{3^{n+[n/2]} \cdot 5831^n \cdot P_n(-1/5831)}{G_n},$$

$$Q_n = \frac{3^{n+[n/2]} \cdot 5831^n \cdot Q_n(-1/5831)}{G_n},$$

and

$$I_n = \frac{3^{n+[n/2]} \cdot 5831^{-n-1} \cdot E_n(-1/5831)}{G_n}.$$

An example (continued)

We thus have a sequence of integers P_n, Q_n with P_n/Q_n and P_{n+1}/Q_{n+1} distinct,

$$|Q_n| < \frac{3^{n+[n/2]} \cdot 5831^n \cdot Q_n(-1/5831)}{G_n} < 42 \times 60598^n$$

and

$$|P_n - \theta Q_n| < \frac{3^{n+[n/2]} \cdot 5831^{-n-1} \cdot E_n(-1/5831)}{G_n} < 43 \times 8975^{-n}.$$

An example (continued)

We thus have a sequence of integers P_n, Q_n with P_n/Q_n and P_{n+1}/Q_{n+1} distinct,

$$|Q_n| < \frac{3^{n+[n/2]} \cdot 5831^n \cdot Q_n(-1/5831)}{G_n} < 42 \times 60598^n$$

and

$$|P_n - \theta Q_n| < \frac{3^{n+[n/2]} \cdot 5831^{-n-1} \cdot E_n(-1/5831)}{G_n} < 43 \times 8975^{-n}.$$

A reminder

If

$$|Q_n| < a \cdot \alpha^n, \quad \text{and} \quad |P_n - Q_n\theta| < b \cdot \beta^{-n},$$

for all positive integers n, then

$$\left|\theta - \frac{p}{q}\right| > \left(2\,a\,\alpha\,(2\,b\,\beta)^\lambda\right)^{-1} |q|^{-1-\lambda} \quad \text{for} \quad \lambda = \frac{\log \alpha}{\log \beta}, \quad (*)$$

for all integers p and $q \neq 0$, with $|q| > 1/2b$.

A reminder

If

$$|Q_n| < a \cdot \alpha^n, \quad \text{and} \quad |P_n - Q_n\theta| < b \cdot \beta^{-n},$$

for all positive integers n, then

$$\left|\theta - \frac{p}{q}\right| > \left(2\,a\,\alpha\,(2\,b\,\beta)^{\lambda}\right)^{-1} |q|^{-1-\lambda} \quad \text{for} \quad \lambda = \frac{\log\alpha}{\log\beta}, \quad (*)$$

for all integers p and $q \neq 0$, with $|q| > 1/2b$.

We have

$$a = 42,\ \alpha = 60598,\ b = 43,\ \beta = 8975$$

An example (concluded)

We thus have

$$\left|\left(1+\frac{1}{5831}\right)^{1/3}-\frac{P}{Q}\right| > \left(7\cdot 10^{13}\right)^{-1} Q^{-2.21},$$

valid for all positive integers P and Q.

An example (concluded)

We thus have

$$\left| \left(1+\frac{1}{5831}\right)^{1/3} - \frac{P}{Q} \right| > \left(7\cdot 10^{13}\right)^{-1} Q^{-2.21},$$

valid for all positive integers P and Q.

But

$$1+\frac{1}{5831} = \frac{5832}{5831} = \frac{18^3}{17\cdot 7^3},$$

and hence writing $P=18q$ and $Q=7p$, we conclude that

$$\left| \frac{18}{7\sqrt[3]{17}} - \frac{18q}{7p} \right| > \left(7\cdot 10^{13}\right)^{-1} (7p)^{-2.21}.$$

An example (concluded)

From the inequality

$$\left| \frac{18}{7\sqrt[3]{17}} - \frac{18q}{7p} \right| > \left(7 \cdot 10^{13}\right)^{-1} (7p)^{-2.21},$$

valid for all positive integers p and q, we conclude that

$$\left| \sqrt[3]{17} - \frac{p}{q} \right| > \left(2 \cdot 10^{16}\right)^{-1} q^{-2.21}$$

and, after a little work, that

$$\left| \sqrt[3]{17} - \frac{p}{q} \right| > \frac{1}{100}\, q^{-2.22}.$$

Solving Nils' problem N1

Suppose we have that

$$x^3 - 17y^3 = 1$$

in, say, positive integers x and y. Then

$$\left|\sqrt[3]{17} - \frac{x}{y}\right| = \frac{1}{x^2y + 17^{1/3}xy^2 + 17^{2/3}y^3} < \frac{1}{3 \cdot 17^{2/3}y^3}.$$

Since also

$$\left|\sqrt[3]{17} - \frac{x}{y}\right| > \frac{1}{100}\, y^{-2.22},$$

it follows that $y < 8$. A quick check yields $(x, y) = (18, 7)$.

More generally

One can prove

Theorem

If a and b are distinct nonzero integers, then the equation

$$\left|ax^3 - by^3\right| = 1$$

has at most one solution in positive integers x and y.

More generally

One can prove

Theorem

If a and b are distinct nonzero integers, then the equation

$$|ax^n - by^n| = 1$$

has at most one solution in positive integers x and y.

A couple more applications

We can use similar arguments to show that

$$\left|x^2 - 2^n\right| > \sqrt{x}$$

for all positive integers x,

$$x \notin \{3, 181, 2^j, j \in \mathbb{Z}\},$$

and to show that writing

$$x^2 + 7 = 2^n \cdot M, \quad \text{for} \;\; M \in \mathbb{Z}$$

implies that either $|x| \in \{1, 3, 5, 11, 181\}$ of $M > \sqrt{|x|}$.