Journal – (2024), —

A Well-Motivated Proof That Pi Is Irrational

Timothy Y. Chow

March 29, 2024

Abstract. Ivan Niven's succinct proof that π is irrational is easy to verify, but it begins with a magical formula that appears to come out of nowhere, and whose origin remains mysterious even after one goes through the proof. The goal of this expository paper is to describe a thought process by which a mathematician might come up with the proof from scratch, without having to be a genius. Compared to previous expositions of Niven's proof, perhaps the main novelty in the present account is an explicit appeal to the theory of orthogonal polynomials, which leads naturally to the consideration of certain integrals whose relevance is otherwise not immediately obvious.

1. Introduction

That π is irrational is something we have all known since childhood, but curiously enough, most mathematicians have either never seen a proof that π is irrational, or have worked through a proof but have found it to be unmotivated and unmemorable. This is not because of a lack of concise proofs; Niven's famous proof ([12] or [2, p. 276]), or a variant thereof [3, 5, 7], occupies less than a page of text, and it is not difficult to check that each step in the proof is correct. A *concise* proof, however, is not the same as what Donald Newman [11] would call a *natural* proof.

> This term ... is introduced to mean not having any ad hoc constructions or *brilliancies*. A "natural" proof, then, is one which proves itself, one available to the "common mathematician in the streets."

Indeed, a verbose proof may be more natural than a concise proof, if the concise proof fails to explain the origin of the underlying ideas. For example, Niven's proof begins by writing down an integral that appears to come out of nowhere. Various authors [8, 10, 15, 16] have attempted to motivate Niven's proof, but I have always been left with the feeling that never in a million years could I have come up with the proof myself.

It was not until recently, after reading Angell's lovely book [1], as well as the answer by Kostya_I [9] to a question that I posted on MathOverflow, that a light bulb went off in my head. The purpose of this paper is to give an account of how Newman's "mathematician in the streets" might discover a proof that π is irrational without having to be a genius.

In order to make this paper as accessible as possible, we do not assume that the reader has any prior familiarity with irrationality proofs. Our discussion proceeds in several stages.

1. We explain the general philosophy behind irrationality proofs, using Fourier's proof that e is irrational as an example.

2. We give a proof that e^r is irrational for positive integers r that minimizes "brilliancies"; ideally, readers will feel that they might have come up with this proof themselves—at least if they were permitted to look up "standard" facts in "standard" references.

3. Using a clever trick, we simplify the proof, to enable the reader to keep it in memory (or reconstruct it) without having to consult any references.

4. Finally, we show that the proof that e^r is irrational for positive integers r can be straightforwardly modified to yield Niven's proof that π is irrational.

2. The fundamental theorem of transcendental number theory

It is an old joke that the fundamental theorem of transcendental number theory is that there is no integer strictly between 0 and 1. Actually, this joke is half serious, because many proofs of irrationality can be framed as follows.

1. Assume toward a contradiction that α is rational. Write down a suitable equation involving α.
2. "Scale up" the equation by a multiple of the denominator of α.
3. Deduce that $A = B$ where A is an integer and $0 < B < 1$.
4. Apply the fundamental theorem of transcendental number theory to deduce a contradiction.

Fourier's proof [14] of the irrationality of e illustrates this schema perfectly. For the purposes of this proof, we define the function e^x in terms of its Taylor series, so

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots \tag{2.1}$$

Assume toward a contradiction that $e = p/q$ for positive integers p and q. Scale up both sides of Equation (2.1) by a factor of $q!$ and observe that some terms are integers; let B denote the sum of the remaining terms.

$$\underbrace{\frac{q!p}{q}}_{\in \mathbb{Z}} = \underbrace{q! + \frac{q!}{1!} + \frac{q!}{2!} + \cdots + \frac{q!}{q!}}_{\in \mathbb{Z}} + \underbrace{\frac{q!}{(q+1)!} + \frac{q!}{(q+2)!} + \cdots}_{B}$$

Since B is the sum of positive terms, $0 < B$. It is straightforward to upper-bound B with a geometric series, which we can sum explicitly.

$$B = \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \cdots < \frac{1}{q+1} + \frac{1}{(q+1)^2} + \cdots = \frac{1}{q} \le 1.$$

This contradicts the fundamental theorem of transcendental number theory; Q.E.D.

It is worth remarking explicitly that what makes this proof work is the *rapid convergence* of Equation (2.1). That is, no matter what p/q we postulate e to be, the "remainder" after scaling up by $q!$ is so small that it lies between 0 and 1.

3. The irrationality of e^r

Emboldened by our success, we might try to adapt Fourier's proof to prove that e^r is irrational when r is a positive integer. Let $e^r = p/q$ as before. Then

$$\frac{p}{q} = 1 + \frac{r}{1!} + \frac{r^2}{2!} + \frac{r^3}{3!} + \cdots$$

What now? The most obvious attempt is to scale up by $q!$ as before.

$$\underbrace{\frac{q!p}{q}}_{\in \mathbb{Z}} = \underbrace{q! + \frac{q!r}{1!} + \cdots + \frac{q!r^q}{q!}}_{\in \mathbb{Z}} + \underbrace{\frac{q!r^{q+1}}{(q+1)!} + \frac{q!r^{q+2}}{(q+2)!} + \cdots}_{B}$$

But now we run into a difficulty. If we try to upper-bound B with a geometric series in the same manner as before, then the bound we get is

$$B < \sum_{n=1}^{\infty} \frac{r^{q+n}}{(q+1)^n} = \frac{r^{q+1}}{(q+1) - r}.$$

We get a contradiction if $r = 1$, but if $r > 1$, then our upper bound on B is large for large q, so our argument fails.

It is quite possible that the above argument can be repaired, but I do not know how to do so. For the purposes of the present discussion, let us assume we get stuck at this point. The Taylor series for e^x is not bringing us joy. What alternatives exist?

4. Orthogonal polynomials to the rescue

We now come to a vital step in the proof. Recall that we desire a rapidly converging approximation to e^x. Where might we find such a thing? The idea is to expand e^x in a basis of *orthogonal polynomials*.

Readers who are not very familiar with orthogonal polynomials will have at least encountered *Chebyshev polynomials of the first kind* (if not by that name) in high-school trigonometry. If we write $\cos n\theta$ as a polynomial in $\cos\theta$, then the polynomials T_n that arise are precisely the Chebyshev polynomials; e.g.,

$$\begin{aligned} T_2(x) &= 2x^2 - 1 \quad \text{because} \quad \cos 2\theta = 2(\cos\theta)^2 - 1, \\ T_3(x) &= 4x^3 - 3x \quad \text{because} \quad \cos 3\theta = 4(\cos\theta)^3 - 3\cos\theta, \\ T_4(x) &= 8x^4 - 8x^2 + 1 \quad \text{because} \quad \cos 4\theta = 8(\cos\theta)^4 - 8(\cos\theta)^2 + 1. \end{aligned}$$

The reason we say that the T_n are orthogonal polynomials is that they turn out to be orthogonal with respect to the following inner product:

$$\langle f, g\rangle := \int_{-1}^{1} f(x)g(x)\,\frac{dx}{\sqrt{1-x^2}}.$$

As in any inner product space, we can expand the vectors (which in this case are functions of x) in terms of a basis of orthogonal vectors, and if the basis is a "good" one, then "nice" functions will be well-approximated.

Once we hit on the idea of computing the coefficients of e^x with respect to a basis of orthogonal polynomials, the question immediately arises: Which set of orthogonal polynomials should we pick? If we are not experts in the subject, we might turn to Wikipedia or some standard reference such as the *NIST Handbook* [13]. We quickly discover a bewildering variety of orthogonal polynomials and associated inner products, such as the following.

$$\begin{aligned} \text{Chebyshev polynomials:} &\quad \langle f, g\rangle := \int_{-1}^{1} f(x)g(x)\,\frac{dx}{\sqrt{1-x^2}} \\ \text{Gegenbauer polynomials:} &\quad \langle f, g\rangle := \int_{-1}^{1} f(x)g(x)(1-x^2)^{\alpha}\,dx \\ \text{Legendre polynomials:} &\quad \langle f, g\rangle := \int_{-1}^{1} f(x)g(x)\,dx \\ \text{Laguerre polynomials:} &\quad \langle f, g\rangle := \int_{0}^{\infty} f(x)g(x)\,e^{-x}\,dx \\ \text{Hermite polynomials:} &\quad \langle f, g\rangle := \int_{-\infty}^{\infty} f(x)g(x)\,e^{-x^2}\,dx \end{aligned}$$

What to do? In the absence of a strong a priori reason to pick one inner product over another, it makes sense to begin with the simplest-looking option—the inner product for Legendre polynomials—and switch horses later if the first attempt does not work. In this way, we arrive at the idea of considering an expression of the form $\int f(x)\,e^x\,dx$ where $f(x)$ is a polynomial.

Before we pull out the *NIST Handbook* to look up standard results on Legendre polynomials, let us note that regardless of which polynomial $f(x)$ we choose, evaluating $\int f(x)\,e^x\,dx$ is likely to require integrating by parts. So let us see what we get if we follow our nose. The parameter r needs to be introduced somehow, so let us change the interval of integration from $[-1,1]$ to $[0,r]$.

$$\begin{aligned}\int_0^r f(x)\,e^x\,dx &= \Big[f(x)\,e^x\Big]_0^r - \int_0^r f'(x)\,e^x\,dx\\ &= \Big[(f(x)-f'(x))\,e^x\Big]_0^r + \int_0^r f''(x)\,e^x\,dx\\ &= \Big[(f(x)-f'(x)+f''(x))\,e^x\Big]_0^r - \int_0^r f'''(x)\,e^x\,dx, \quad \text{etc.}\end{aligned}$$

Therefore, if we define $F(x) := f(x)-f'(x)+f''(x)-f'''(x)+\cdots$ (there are no convergence problems, because $f(x)$ is a polynomial), then

$$\int_0^r f(x)\,e^x\,dx = F(r)\,e^r - F(0). \tag{4.2}$$

If the fundamental theorem of transcendental number theory is burned into our minds, Equation (4.2) should make us perk up. Suppose toward a contradiction that $e^r = p/q$. Can we choose $f(x)$ so that, after scaling up by q, the right-hand side is an integer, but the left-hand side is strictly between 0 and 1? There seems to be a tension; the simplest way to ensure that the right-hand side of Equation (4.2) is an integer (after multiplying by q) is to require $f(x)$ to have integer coefficients, because then $F(r)$ and $F(0)$ will be integers. But if $f(x)$ has integer coefficients, then there would appear to be no reason for the integral on the left-hand side to be small.

5. Legendre polynomials

It is time to look up the theory of Legendre polynomials, to see if it can help us. The usual definition of Legendre polynomials assumes an interval of integration of $[-1,1]$; one definition [13, 18.5.8] is

$$P_n(x) = \frac{1}{2^n}\sum_{\ell=0}^{n}\binom{n}{\ell}^2 (x-1)^{n-\ell}(x+1)^\ell.$$

Recall that we have changed the interval of integration to $[0,r]$, so we need to make an affine change of variables. Define

$$\tilde{P}_n(x) := P_n(2x/r-1) = \frac{1}{r^n}\sum_{\ell=0}^{n}\binom{n}{\ell}^2 x^\ell (x-r)^{n-\ell}. \tag{5.3}$$

Now, the nth coefficient in the Legendre polynomial expansion of a function such as e^x is (up to a normalizing factor that we will ignore for now) obtained by taking the inner product of e^x with the nth Legendre polynomial:

$$\int_0^r \tilde{P}_n(x)\,e^x\,dx.$$

As we alluded to earlier, it is natural to guess that Legendre polynomials form a "good" basis in the sense that these coefficients shrink rapidly as n increases. Aha! We want the integral in Equation (4.2) to be small, so maybe setting $f(x) := \tilde{P}_n(x)$ is the key? But remember that we also want $f(x)$ to have integer coefficients. Equation (5.3) implies that $r^n\tilde{P}_n(x)$ has integer coefficients, so let us try setting $f(x) := r^n\tilde{P}_n(x)$. Then Equation (4.2) becomes (after multiplying by q)

$$r^n q\int_0^r \tilde{P}_n(x)\,e^x\,dx = F(r)\cdot p - F(0)\cdot q,$$

where $F(x) = f(x) - f'(x) + f''(x) - \cdots$. The right-hand side is an integer, so to obtain our desired contradiction, it would be enough if

$$r^n q \int_0^r \tilde{P}_n(x)\, e^x \, dx \tag{5.4}$$

were nonzero, with absolute value less than 1 for some n. But is this the case?

It turns out that the answer is yes! The proof is a relatively routine but somewhat tedious calculation that we will present in a moment. We therefore have:

Theorem 1. *If r is a positive integer, then e^r is irrational.*

Let us pause a moment to take stock of the situation. We claim that we have managed to find a proof that e^r is irrational (for positive integers r) with a minimum of flashes of brilliance. Once we think to use orthogonal polynomials—and in particular Legendre polynomials—to give us a good approximation to e^x, we are led to consider an integral (over a finite interval) of the form $\int f(x)\, e^x\, dx$ where $f(x)$ is a (suitably normalized) Legendre polynomial. On the one hand, we can calculate directly that the integral, even after multiplication by the denominator of e^r, shrinks rapidly. On the other hand, we can use integration by parts to show that its value is an integer, contradicting the fundamental theorem of transcendental number theory.

First proof of Theorem 1. The following argument is not needed for anything else in this paper, so the reader who is unfamiliar with orthogonal polynomials is encouraged to skip ahead to Section 6. and return here later.

As we said above, we just need to prove that (5.4) is nonzero and has absolute value less than 1 for some n. The reason (5.4) is nonzero for infinitely many n is that Legendre polynomials form an orthogonal basis for the Hilbert space of square-integrable functions on a compact interval, so having zero coefficients for all sufficiently large n would imply that e^x is a *finite* sum of polynomials, which of course is absurd.

The polynomial $\tilde{P}_n(x)$ is orthogonal to all polynomials of degree less than n, so the value of the integral (5.4) remains unchanged if we replace e^x by $e_n(x)$, which we define to be e^x minus the terms in its Taylor series of degree less than n. The Cauchy–Bunyakovsky–Schwarz inequality implies that

$$\begin{aligned}
\left|\int_0^r \tilde{P}_n(x)\, e^x\, dx\right| &= \left|\int_0^r \tilde{P}_n(x)\, e_n(x)\, dx\right| \\
&\le \int_0^r \left|\tilde{P}_n(x)\, e_n(x)\right| dx \\
&\le ||\tilde{P}_n(x)||_2 \cdot ||e_n(x)||_2,
\end{aligned}$$

where $||\cdot||_2$ denotes the $L^2[0,r]$ norm. Using the substitution $y = 2x/r - 1$,

$$||\tilde{P}_n(x)||_2^2 = \int_0^r \tilde{P}_n(x)^2\, dx = \frac{r}{2}\int_{-1}^1 P_n(y)^2\, dy = \frac{r}{2n+1},$$

where the last equality above is a standard fact [13, Table 18.3.1] about the squared norm of the Legendre polynomial. As for $e_n(x)$, for $0 \le x \le r$ and large n,

$$\begin{aligned}
e_n(x) &= \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!} + \cdots \le \frac{x^n}{n!}\left(1 + \frac{x}{n+1} + \frac{x^2}{(n+1)^2} + \cdots\right) \\
&= \frac{x^n}{n!(1 - x/(n+1))} \le \frac{2x^n}{n!}.
\end{aligned}$$

Therefore

$$||e_n(x)||_2 = \sqrt{\int_0^r e_n(x)^2\,dx} \le \frac{2}{n!}\sqrt{\int_0^r x^{2n}\,dx}$$
$$= \frac{2r^{(2n+1)/2}}{n!\sqrt{2n+1}}.$$

Putting this all together yields the following upper bound on the absolute value of (5.4):

$$r^n q \cdot \sqrt{\frac{r}{2n+1}} \cdot \frac{2r^{(2n+1)/2}}{n!\sqrt{2n+1}},$$

which for fixed r and large n is less than 1, because exponential functions grow much more slowly than factorials. □

6. A shorter and more elementary proof

Readers who are satisfied that the proof we have just presented is well-motivated may nevertheless be dissatisfied that it requires us to know (or re-derive) various facts about Legendre polynomials, and that the calculations are a bit tedious. The conceptual outline of the proof is not hard to remember, and it motivates why we want to consider an integral of the form $\int f(x)\,e^x\,dx$, but the details are not so memorable.

In this section, we rectify these defects with an alternative version of the proof that is simpler and requires no "theory." The price we pay is that we will need to invoke a clever idea.

Let us revisit Equation (4.2). Assuming $e^r = p/q$ and $f(x)$ is a polynomial,

$$q\int_0^r f(x)\,e^x\,dx = F(r)\cdot p - F(0)\cdot q. \tag{6.5}$$

where $F(x) := f(x) - f'(x) + f''(x) - f'''(x) + \cdots$. Recall that we want to choose $f(x)$ so that the left-hand side of Equation (6.5) is small (between 0 and 1), and at the same time the right-hand side is an integer. Can we figure out from first principles how to choose a suitable $f(x)$?

Here is the clever idea: the definition of $F(x)$ involves high-order derivatives $f^{(n)}(x)$, and *repeated differentiation of a polynomial produces a factorial-like coefficient out in front!* That suggests that if $f(x)$ has integer coefficients, then not only will $f^{(n)}(x)$ have integer coefficients, but it will still have integer coefficients even after dividing by $n!$. Indeed:

Lemma 1. *If $f(x)$ is a polynomial with integer coefficients, then for any nonnegative integer n, $f^{(n)}(x)/n!$ has integer coefficients.*

Proof. It suffices to prove Lemma 1 in the case $f(x) = x^m$. The nth derivative of x^m is $m(m-1)\cdots(m-n+1)x^{m-n}$, and dividing the coefficient by $n!$ yields the binomial coefficient $\binom{m}{n}$, which is an integer for any nonnegative integer m. □

Lemma 1 is promising, because dividing by $n!$ should make the left-hand side of Equation (6.5) small. But there is still a difficulty; if $k \ge n$, then we can divide $f^{(k)}(x)$ by $n!$ and still get integer coefficients, but what do we do about those pesky lower-order derivatives ($k < n$) that show up in $F(x)$? We solve this problem by taking advantage of our freedom to choose $f(x)$. If $f(x)$ vanishes to order n at some point $x = c$, then $f^{(k)}(c)$ vanishes for $k < n$. Since the right-hand side of Equation (6.5) involves $F(0)$ and $F(r)$, we are motivated to consider the polynomial $x^n(r-x)^n$ (or the polynomial $x^n(x-r)^n$, but we will see soon why $x^n(r-x)^n$ better serves our purposes), which vanishes to order n at both $x = 0$ and $x = r$.

Corollary 1. *For integers r and n with $n \geq 0$, let $f(x) = x^n(r-x)^n$, and $F(x) = f(x) - f'(x) + f''(x) - \cdots$. Then $F(0)/n!$ and $F(r)/n!$ are integers.*

Proof. Since $f(x)$ vanishes to order n at $x = 0$ and $x = r$, $f^{(k)}(0) = f^{(k)}(r) = 0$ if $k < n$. If $k \geq n$, then Lemma 1 tells us that $f^{(k)}(x)/n!$ has integer coefficients. Thus each summand of $F(0)/n!$ and $F(r)/n!$ is either zero or an evaluation of a polynomial with integer coefficients at an integer value. □

We are now in a position to give a second proof of Theorem 1 that does not require any knowledge of orthogonal polynomials.

Second proof of Theorem 1. Let $f(x) := x^n(r-x)^n$ for some n (whose value will be chosen later), and define $F(x) := f(x) - f'(x) + f''(x) - f'''(x) + \cdots$. Assume toward a contradiction that $e^r = p/q$ for positive integers p and q. Equation (6.5) implies

$$\frac{q}{n!}\int_0^r x^n(r-x)^n\, e^x\, dx = \frac{F(r)p}{n!} - \frac{F(0)q}{n!}. \tag{6.6}$$

Corollary 1 tells us that the right-hand side of Equation (6.6) is an integer. On the left-hand side, the interval of integration has nonzero length because $r > 0$, and the integrand is strictly positive between 0 and r (this is where our choice of $x^n(r-x)^n$ instead of $x^n(x-r)^n$ makes a difference), so the left-hand side is strictly positive. But we can also upper-bound the integral by multiplying the length of the interval of integration by a term-by-term upper bound on the factors of the integrand:

$$\int_0^r x^n(r-x)^n\, e^x\, dx \leq r \cdot r^n \cdot r^n \cdot e^r.$$

Because $n!$ grows super-exponentially fast, the left-hand side of Equation (6.6) is less than 1 for all sufficiently large n, contradicting the fundamental theorem of transcendental number theory. □

We remark in passing that the two proofs of Theorem 1 are not as different as they might seem at first glance, because the Rodrigues formula [13, 18.5.5] for Legendre polynomials implies

$$\tilde{P}_n(x) = \frac{1}{n!r^n}\frac{d^n}{dx^n}(x^n(x-r)^n),$$

so $\tilde{P}_n(x)$ is in fact closely related to the polynomial $f(x) = x^n(r-x)^n$.

7. On to the irrationality of π

We have worked rather hard, and have not even mentioned π yet. But as we promised earlier, a small modification of the ideas we have already seen will prove the irrationality of π.

First, though, we should at least briefly address the basic question, just what *is* π anyway? The least controversial definition is that π is the limit of the perimeter of a regular n-gon of unit diameter as n goes to infinity. Unfortunately, this geometric definition is awkward to work with. Instead, what we do is to define

$$\sin x := x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

and we define π to be the smallest positive number such that $\sin \pi = 0$. The advantage of this approach is that it allows us to apply the tools of calculus to study π. Now, one could object that it is far from obvious that this definition of π coincides with the goemetric one. This objection is valid, but we take the view that the equivalence of the two definitions is a standard fact that the "common mathematician in the streets" would be familiar with.

Theorem 1 says that e^r is irrational for all positive integers r, but in fact it immediately implies that e^r is irrational for all nonzero rational r, because if e^r is irrational then so is $e^{-r} = 1/e^r$, and if

$e^{r/a}$ is rational then so is $(e^{r/a})^a = e^r$. We can restate this result in the following form: if r is nonzero and e^r is rational, then r is irrational. To prove that π is irrational, *we run through almost the same argument with the function* $\sin x$ *instead of the function* e^x*, and we take* $r = \pi$. Now $\sin r = \sin \pi = 0$ is manifestly rational, and we will argue that this implies that π is irrational.

There are only two adjustments that need to be made in the proof. First, replacing e^x with $\sin x$ changes the integration by parts formula slightly.

$$\begin{aligned}\int_0^\pi f(x)\,\sin x\,dx &= \Big[-f(x)\,\cos x\Big]_0^\pi + \int_0^\pi f'(x)\,\cos x\,dx \\ &= f(\pi) + f(0) + \Big[f'(x)\,\sin x\Big]_0^\pi - \int_0^\pi f''(x)\,\sin x\,dx \\ &= f(\pi) + f(0) - \int_0^\pi f''(x)\,\sin x\,dx.\end{aligned}$$

So if we define $F(x) := f(x) - f''(x) + f''''(x) - f''''''(x) + \cdots$ then

$$\int_0^\pi f(x)\,\sin x\,dx = F(\pi) + F(0). \tag{7.7}$$

The second adjustment arises because we will be assuming toward a contradiction that $\pi = a/b$, and since π is playing the role that r was playing, we need a version of Corollary 1 in which $r - x$ is replaced with $a - bx$.

Corollary 2. *For integers* a, b, n *with* $b \neq 0$ *and* $n \geq 0$*, let* $f(x) = x^n(a - bx)^n$*, and* $F(x) = f(x) - f''(x) + f''''(x) - \cdots$. *Then* $F(0)/n!$ *and* $b^n F(a/b)/n!$ *are integers.*

Proof. As in the proof of Corollary 1, since $f(x)$ vanishes to order n at $x = 0$ and $x = a/b$, $f^{(k)}(0) = f^{(k)}(a/b) = 0$ if $k < n$. If $k \geq n$, then Lemma 1 tells us that $f^{(k)}(x)/n!$ has integer coefficients. Each summand of $F(0)/n!$ is either zero or the constant term of a polynomial with integer coefficients, so $F(0)/n!$ is an integer. As for $F(a/b)/n!$, it is the result of evaluating a degree-n polynomial in x with integer coefficients at $x = a/b$, so multiplying it by b^n yields an integer. Q.E.D.

Remark. In fact, something even stronger is true; the hypotheses of Corollary 2 imply that $F(a/b)/n!$ is already an integer, because n-fold differentiation of $f(x)$ produces suitable powers of b via the chain rule. But we do not need this stronger result.

Theorem 2. π *is irrational.*

Proof. Assume toward a contradiction that $\pi = a/b$ for positive integers a and b. We define the polynomial $f(x) := x^n(a-bx)^n$, and set $F(x) := f(x) - f''(x) + f''''(x) - f''''''(x) + \cdots$. By Equation (7.7),

$$\int_0^\pi f(x)\,\sin x\,dx = F(\pi) + F(0).$$

If we multiply both sides of the equation by $b^n/n!$, then we obtain

$$\frac{b^n}{n!}\int_0^{a/b} x^n(a - bx)^n\,\sin x\,dx = \frac{b^n F(a/b)}{n!} + \frac{b^n F(0)}{n!}. \tag{7.8}$$

Corollary 2 implies that the right-hand side of Equation (7.8) is an integer. On the left-hand side, the integrand is strictly positive between 0 and a/b (this is where our decision to use the function $\sin x$, rather than $\cos x$ or some other phase shift, makes a difference), and the interval of integration is nonzero because $\pi > 0$, so the left-hand side is strictly positive. But we can also upper-bound the integral by multiplying the length of the interval of integration by a term-by-term upper bound on the factors of the integrand:

$$\int_0^{a/b} x^n(a-bx)^n \sin x\, dx \le \frac{a}{b}\left(\frac{a}{b}\right)^n a^n.$$

Because $n!$ grows super-exponentially fast, the left-hand side of Equation (7.8) is less than 1 for all sufficiently large n, contradicting the fundamental theorem of transcendental number theory. □

We should remark that it is not hard to make similar adjustments to the first proof of Theorem 1, in Section 5., to prove the irrationality of π (see [9] for details). Some readers may prefer this proof, if the "clever idea" of Section 6. seems too ad hoc.

8. Cheat sheet

Despite the brevity of Niven's proof, I was never previously able to retain it in my long-term memory, because I did not understand where all the ingredients came from. But now, I am confident that I could reconstruct it without having to look anything up. I have found that the following "cheat sheet" is all that I need to remember, and I offer it to the reader as a mnemonic aid.

1. Motivated by the theory of orthogonal polynomials, we consider an expression of the form $\int f(x) \sin x\, dx$ where $f(x)$ is a suitably chosen polynomial.
2. Integration by parts implies that if $F(x) := f(x) - f''(x) + f''''(x) - \cdots$ then
$$\int_0^{\pi} f(x) \sin x\, dx = F(\pi) + F(0).$$
3. If $\pi = a/b$, then we want the right-hand side (after suitable scaling) to be an integer, and the left-hand side to be between 0 and 1. We do this by choosing f to vanish to high order n at 0 and a/b, and then multiplying both sides by $b^n/n!$. The denominator of $n!$ makes the LHS small. The high-order vanishing implies that the RHS is an integer, because low-order derivatives of f vanish at 0 and π, and high-order derivatives of f have coefficients that are divisible by $n!$.

9. Historical remarks and acknowledgments

As we mentioned earlier, the irrationality proof of π that we present here is due to Niven [12], although the ideas can be traced back to Hermite [6] and [2, pp. 162–193]. Our exposition, except for the connection to orthogonal polynomials, borrows heavily from Angell [1]. The proof using Legendre polynomials is adapted from the answer by Kostya_I [9] to a question that I posted on MathOverflow, asking about ways to motivate the proof that π is irrational. Kostya_I's proof was a revelation to me, and a major inspiration for my writing this article in the first place.

The educated reader will notice that this paper conspicuously fails to mention continued fractions, which play an important role in the study of rational approximations of irrational numbers. Continued fractions were used by Lambert [2, pp. 141–146] in the first ever proof that π is irrational, and also show up in Hermite's work in the form of what we now call *Padé approximants* (although Hermite came before Padé). Padé approximants in turn are closely related to orthogonal polynomials, so the relevance of orthogonal polynomials to the irrationality of π is not unexpected, but we can perhaps claim some novelty in the way we have emphasized their usefulness for expository purposes. For more discussion of the relationship between Equation (4.2) and Padé approximants and the continued fraction expansion of e, see Cohn's *Monthly* note [4].

We have emphasized the parallel between the irrationality of e^r and the irrationality of π, but we should point out that the seemingly innocuous contraposition when switching from "r rational implies e^r irrational" to "$\sin\pi$ rational implies π irrational" means that the proof technique yields explicit rational approximations to e^r, but does *not* yield explicit rational approximations to π. Finding high-quality explicit rational approximations to π is a topic that is beyond the scope of this article.

Finally, we thank Nicholas Nguyen, John Zuehlke, and especially Peter Farbman for carefully reading drafts of this paper; their input has significantly improved the exposition.

References

[1] Angell, D., *Irrationality and Transcendence in Number Theory*, CRC Press, 2022.

[2] Berggren, J. L., Borwein, J., and Borwein, P., *Pi: A Source Book*, 3rd edition, Springer, 2014.

[3] Bourbaki, N., *Fonctions d'une Variable Réelle, Chapitres 1, 2, et 3*, Actualités Scientifiques et Industrielles, vol. 1074, Hermann, 1949, pp. 137–138.

[4] Cohn, H., A short proof of the simple continued fraction expansion of e, *Amer. Math. Monthly* **113** (2006), 57–62.

[5] Hardy, G. H., and Wright, E. M., *Introduction to the Theory of Numbers*, 6th edition, Oxford University Press, 2008, p. 55.

[6] Hermite, C., Sur quelques approximations algébriques, *J. de Crelle* **76** (1873), 342–344.

[7] Jeffreys, H., *Scientific Inference*, 3rd edition, Cambridge University Press, 1973, p. 268.

[8] Jones, T. W., Discovering and proving that π is irrational, *Amer. Math. Monthly* **117** (2010), 553–557.

[9] Kostya_I, To prove irrationality, why integrate?, `https://mathoverflow.net/a/444167` (version: 2023-04-14).

[10] Müller, J., and Müller, T., 3 in 1: A simple way to prove that e^r, $\ln(r)$ and π^2 are irrational, *Elem. Math.* **73** (2018), 133–135.

[11] Newman, D., *Analytic Number Theory*, Springer-Verlag, 1998, p. 59.

[12] Niven, I., A simple proof that π is irrational, *Bull. Amer. Math. Soc.* **53** (1947), 509.

[13] Olver, F. W. J. et al. (eds.), *NIST Handbook of Mathematical Functions*, Cambridge University Press, 2010.

[14] de Stainville, J., *Mélange d'analyse algébrique et de géométrie*, Veuve Courcier, 1815, pp. 340–341.

[15] Zhou, L., Irrationality proofs à la Hermite, *Math. Gazette* **95** (2011), 407–413.

[16] Zhou, L.; On "Discovering and proving that π is irrational," *Mitt. Dtsch. Math.-Ver.* online archive (2011), `https://page.math.tu-berlin.de/~mdmv/archive/19/mdmv-19-zhou.pdf`