**Parading a panoply of prime proofs.**

# A panopoly of proofs that there are infinitely many primes

**Andrew Granville**

Abstract: There are many different ways to prove that there are infinitely many primes. I will highlight a few of my favourites, selected so as to involve a rich variety of mathematical ideas.

## 1 Different types of proofs

### 1.1 Proofs by contradiction

Proofs that there are infinitely many primes typically rely on the theorem that

*Every integer $q > 1$ has a prime factor.*

Euclid used this to prove that there are infinitely many primes, as follows: Suppose that $p_1, \dots, p_k$ is a complete list of all of the primes. Now $q = p_1 \cdots p_k + 1$ is divisible by some prime $p$. But then $p = p_j$ for some $j$ and so $q \equiv 1 \pmod{p}$, so that $(q,p) = 1$, a contradiction. □

There are many variations on this theme. For instance we can take $q$ to be $p_1 \cdots p_k - 1$, or $mp_1 \cdots p_k + 1$ for any integer $m \neq 0$. We could also split the primes into any two subsets: write $\{p_1, \dots, p_k\} = \mathcal{M} \cup \mathcal{N}$, and then let $m$ be the product of the elements of $\mathcal{M}$, and let $n$ be the product of the elements of $\mathcal{N}$. Finally let $q = m+n$ have prime divisor $p$. Then $p$ must divide one, and only one, of $m$ and $n$: if $p$ divides, say, $m$ then $(q,p) = (n,p) = 1$, a contradiction. □

One could also take $q = |m-n|$, and as long as this is not $1$ then the analogous argument works, though there are a couple of examples known where $m - n = 1$.[1]

One can have more than two summands: If $N = p_1 \cdots p_k$, let $q = \sum_{i=1}^k N/p_i$. Now $p_j$ divides $N/p_i$ whenever $i \neq j$, so that $(q, p_j) = (N/p_j, p_j) = 1$. A more flexible variant comes by including coefficients $c_1, \dots, c_k$ where each $c_j$ is an integer that is not divisible by $p_j$, and then $q = \sum_{i=1}^k c_i \ N/p_i$. This is so flexible that if $p_1, \dots, p_k$ are the primes up to $x$, then each prime between $x$ and $x^2$ equals such a $q$, for carefully selected values of the $c_j$ (see [1]).

The key idea in Euclid's proof is that $q$ is an integer that is greater than $1$ and coprime to $N = p_1 \cdots p_k$. This can easily be generalized as Euler showed that there are $(p_1 - 1)(p_2 - 1) \cdots (p_k - 1)$ positive integers $\leq N$ that are coprime to $N$, so we could have taken $q$ to be any such integer $> 1$.

### 1.2 A (point-set) topological proof

One of the most elegant ways to present Euclid's idea is in Furstenberg's extraordinary proof [5] using basic notions of point set topology:

Define a topology on the set of integers $\mathbb{Z}$ in which a set $S$ is open if it is empty or if for every $a \in S$ there is an arithmetic progression

$$\mathbb{Z}(a,m) := \{a + nm : n \in \mathbb{Z}\},$$

with $m \neq 0$, which is a subset of $S$. Evidently each $\mathbb{Z}(a,m)$ is open, and it is also closed since

$$\mathbb{Z}(a,m) = \mathbb{Z} \setminus \bigcup_{b:\ 0 \leq b \leq m-1,\ b \neq a} \mathbb{Z}(b,m).$$

If there are only finitely many primes $p$ then $A = \cup_p \ \mathbb{Z}(0,p)$ is also closed, and so $\mathbb{Z} \setminus A = \{-1, 1\}$ is open, but this is false since $\{-1,1\}$ is finite and so cannot contain any arithmetic progression $\mathbb{Z}(a,m)$, as this would contain infinitely many integers. This contradiction implies that there are infinitely many primes. □

I love the surprising sparse elegance of this proof. However, I know of other number theorists who dislike the way it obscures what is really going on.

### 1.3 An analytic proof

The idea is to count the number of positive integers up to some large point $x$ whose prime factors only come from a given set of primes $\mathcal{P} = \{p_1 < p_2 < \ldots < p_k\}$. These integers all take the form

$$p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k} \text{ for some integers } e_j, \text{ each } \geq 0. \tag{1}$$

We are going to count the number of such integers up to $x = 2^m - 1$, for an arbitrary integer $m \geq 1$, by studying this formula: For each $j$, the prime $p_j \geq 2$, and every other $p_i^{e_i} \geq 1$, and so

$$2^{e_j} \leq p_j^{e_j} \leq p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k} \leq 2^m - 1.$$

This implies that $e_j$ is at most $m - 1$, and so there are at most $m$ possibilities for the integer $e_j$, the integers from $0$ through to $m - 1$. Therefore the

[1] Most famously, at least for baseball afficionados, Babe Ruth's home runs record of $714 = 2 \times 3 \times 7 \times 17$ home runs, was overtaken when Hank Aaron hit $715 = 5 \times 11 \times 13$.

number of integers of the form (1), up to $2^m - 1$, is

$$\leq \prod_{j=1}^{k} \#\{\text{integers } e_j : 0 \leq e_j \leq m-1\} = m^k.$$

Now if $\mathcal{P}$ is the set of all primes then every positive integer is of the form (1), and so the last equation implies that $2^m - 1 \leq m^k$ for all integers $m$. We select $m = 2^k$, so this implies that $2^k \leq k^2$, which is false for every integer $k \geq 5$. Therefore as we know that there are at least five primes (for example $2, 3, 5, 7, 11$), we can deduce that there cannot be finitely many. □

This proof highlights the use of counting arguments in number theory, a first step on the road to analytic number theory.

## 1.4 Two arithmetic proofs.

Fermat's little theorem implies that if $p$ is an odd prime then

$$2^{p-1} \equiv 1 \pmod{p}.$$

If $2^m \equiv 1 \pmod{p}$ then one can deduce that $2^g \equiv 1 \pmod{p}$ where $g = (m, p-1)$. We will use this observation to give two proof of the infinitude of primes, booth based on arithmetic structure.

- Suppose that there are only finitely many primes and let $q$ be the largest prime. If $p$ is a prime factor of the Mersenne number, $2^q - 1$, then $2^q \equiv 1 \pmod{p}$. Therefore $2^g \equiv 1 \pmod{p}$ where $g = \gcd(q, p-1)$. Now $g$ divides $q$, so $g$ must equal either $1$ or $q$. However $g$ cannot equal 1, else $p$ divides $2^g - 1 = 2 - 1 = 1$. Therefore $q = g$ which divides $p - 1$. But then $q \leq p - 1 < p$, so $p$ is a larger prime than $q$, contradicting the maximality of $q$. □

- Suppose that there are only finitely many primes $p_1, \ldots, p_k$ and let $2^n$ be the highest power of $2$ dividing any of the $p_j - 1$. Let $q = 2^{2^n} + 1$, the $n$th Fermat number. Then $2^{2^{n+1}} - 1 = (2^{2^n})^2 - 1 \equiv (-1)^2 - 1 \equiv 0 \pmod{q}$, and so if $p$ is a prime factor of $q$ then $2^{2^{n+1}} \equiv 1 \pmod{p}$. Therefore $2^g \equiv 1 \pmod{p}$ where $g = \gcd(2^{n+1}, p-1)$. Now $g$ divides $2^{n+1}$ so $g$ must be a power of 2, say $g = 2^m$. Moreover $m \leq n$ as $g = 2^m$ divides $p - 1$, and $2^n$ was defined to be the highest power of $2$ dividing any $p_j - 1$. Therefore

$$\begin{aligned} 0 \equiv q = 2^{2^n} + 1 &= (2^{2^m})^{2^{n-m}} + 1 \\ &\equiv 1^{2^{n-m}} + 1 \equiv 2 \pmod{p}, \end{aligned}$$

so that $p$ divides $2$, which is impossible as $p$ is an odd prime. □

This proof also yields that for any integer $N \geq 1$, there are infinitely many primes $\equiv 1 \pmod{2^N}$, and suitable modifications even allow one to prove that for any integer $m \geq 2$, there are infinitely many primes $\equiv 1 \pmod{m}$. In 1837 Dirichlet proved that if $(a, q) = 1$ then there are infinitely many primes $\equiv a \pmod{q}$. Far ahead of his time, Dirichlet used analytic methods to prove this result. There is still no known elementary proof of this fact for all pairwise coprime $a$ and $q$, though here we have indicated an approach that works whenever $a = 1$.

## 1.5 A proof by irrationality

Euler exhibited the inspiring identity

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \ldots$$

Let $\delta(n) = 1$ or $-1$ as $n \equiv 1$ or $-1 \pmod{4}$, and $\delta(n) = 0$ if $n \equiv 0 \pmod{2}$. The key observation is that if an odd integer $n$ factors as in (1), then $\delta(n)/n = \prod_{j=1}^{k} (\delta(p_j)/p_j)^{e_j}$. Therefore if there are only finitely many primes then the right-hand side of Euler's identity can be separated into the contributions from each prime, to obtain the identity,

$$\frac{\pi}{4} = \prod_{\substack{p \text{ prime} \\ p \equiv 1 \pmod{4}}} \frac{p}{p-1} \cdot \prod_{\substack{p \text{ prime} \\ p \equiv 3 \pmod{4}}} \frac{p}{p+1}.$$

It is well-known that $\pi$ (and so $\pi/4$) is irrational, but under the assumption that there are only finitely many primes, the right-hand side is a finite product of rational numbers, so is rational, a contradiction. □

The function $\delta$ is periodic of period $4$, equals $0$ whenever $(n, 4) > 1$, and otherwise equals $1$ or $-1$, sums over its period to $0$ (as $1 + 0 + (-1) + 0 = 0$), and factors much like the integers, in that $\delta(n) = \prod_{j=1}^{k} \delta(p_j)^{e_j}$. For every integer $m > 2$ with $m \not\equiv 2 \pmod{4}$, there exists such a function $\delta$ with "4" replaced by "$m$" in the definition. The sum of Euler's series, $\sum \delta(n)/n$ is the "special value" of Dirichlet's $L$-function that is central to his proof that there are infinitely many primes in arithmetic progressions. Moreover, much like here, if $m = 4k$ where $k$ is not divisible by any squares and $k \equiv 1 \pmod{4}$, then the sum adds up to a rational multiple of $\pi\sqrt{|m|}$.

Euler's work pre-dated Dirichlet by almost 100 years, yet he developed the theory of this same mathematical construction without knowing how important it would become. Such prescience can be found in the works of great mathematicians.

## 1.6 A proof by combinatorics and arithmetic geometry

Van der Waerden's Theorem, a deep result in combinatorics, states that for any given $m \geq 2$ and $\ell \geq 3$, if every positive integer is assigned one of $m$ colours, in any way at all, then there is an $\ell$-term arithmetic progression of integers which each have the same colour. Alpoge [3], a current Ph.D. student at Harvard, suggested the following clever colouring of the integers, assuming that $p_1, \ldots, p_k$ are all the primes. Each integer $n$ factors as in (1); we write each exponent $e_j \equiv r_j \pmod 2$ with $r_j = 0$ or 1. Writing $R = p_1^{r_1} \cdots p_k^{r_k}$ we note that $n/R$ is the square of an integer, and we "colour" $n$ with the colour $R$. There are $2^k$ possibilities for $R$. By applying van der Waerden's Theorem with $m = 2^k$ and $\ell = 4$ we deduce that there are four integers in arithmetic progression

$$A, A+D, A+2D, A+3D, \text{ with } D \geq 1,$$

which all have the same colour $R$. Now $R$ divides each of these numbers, so also divides $D = (A+D) - A$. Letting $a = A/R$ and $d = D/R$, we deduce that

$$a, a+d, a+2d, a+3d$$

are four squares in arithmetic progression ($a + jd = (A + jD)/R$ is a square as $A + jD$ has colour $R$.) However Fermat proved (and this is often covered in a first course on elliptic curves) that there cannot be four squares in an arithmetic progression. □

Although this proof uses two far deeper theorems than Euclid's original proof, one cannot help but be charmed by how they can be combined in this way. Actually these ideas have come together before in an unlikely way ([4]), in bounding the number of squares that can possibly appear in an $N$-term arithmetic progression.

## 1.7 The construction of infinitely many primes.

We want to construct an infinite sequence of distinct, pairwise coprime, integers $a_0, a_1, \ldots$; that is, a sequence for which $\gcd(a_m, a_n) = 1$ whenever $m \neq n$. Let $p_n$ be a prime divisor of $a_n$ whenever $|a_n| > 1$. Then the $p_n$ form an infinite sequence of distinct primes. (For, if not, then $p_m = p_n$ for some $m \neq n$ and so $p_m = (p_m, p_n)$ divides $(a_m, a_n) = 1$, a contradiction.) Here a couple of ways to construct such sequences:

- Modification of Euclid's proof: Let $a_0 = 2, a_1 = 3$ and

$$a_n = a_0 a_1 \cdots a_{n-1} + 1 \text{ for each } n \geq 1.$$

If $m < n$ then $a_m$ divides $a_0 a_1 \ldots a_{n-1} = a_n - 1$ and so $\gcd(a_m, a_n)$ divides $\gcd(a_n - 1, a_n) = 1$, which implies that $\gcd(a_m, a_n) = 1$. Therefore if $p_n$ is a prime divisor of $a_n$ for each $n \geq 0$, then $p_0, p_1, \ldots$ is an infinite sequence of distinct primes.

The recurrence for the $a_n$ can be re-written as

$$\begin{aligned} a_{n+1} &= a_0 a_1 \cdots a_{n-1} \cdot a_n + 1 \\ &= (a_n - 1)a_n + 1 = f(a_n), \end{aligned}$$

where $f(x) = x^2 - x + 1$, which leads us to a different proof that these numbers are pairwise coprime. We use the fact that for any distinct integers $r$ and $s$, and any polynomial $f(x) \in \mathbb{Z}[x]$, $r - s$ always divides $f(r) - f(s)$. Therefore if $r \equiv s \pmod p$ then $f(r) \equiv f(s) \pmod p$. Therefore if $p$ divides $a_n$ then $a_{n+1} = f(a_n) \equiv f(0) = 1 \pmod p$. Next $a_{n+2} = f(a_{n+1}) \equiv f(1) = 1 \pmod p$, and then $a_{n+3} = f(a_{n+2}) \equiv f(1) = 1 \pmod p$, and proceeding like this,

$$\begin{aligned} a_{n+k} &= f(f(f(\ldots f(a_{n+1}) \ldots))) \\ &\equiv f(f(f(\ldots f(1) \ldots))) \equiv 1 \pmod p, \end{aligned}$$

for all $k \geq 1$; and so we deduce that $a_m \equiv 1 \pmod p$ for all $m > n$. We deduce that $a_m$ and $a_n$ cannot share any prime factor $p$, and so are coprime. □

- Fermat claimed that the integers $F_n = 2^{2^n} + 1$ are primes for all $n \geq 0$. This is true for $3, 5, 17, 257, 65537$, but false for $F_5 = 641 \times 6700417$, as noted by Euler.[2] Nonetheless the $F_n$ are pairwise coprime, and so we can deduce that if $p_n$ is a prime divisor of $F_n$, then $p_0, p_1, \ldots$ is an infinite sequence of distinct primes. To prove this we begin by noting that the $F_n$-values can be determined by a simple recurrence, as follows:

$$\begin{aligned} F_{n+1} &= (2^{2^n} + 1)(2^{2^n} - 1) + 2 \\ &= F_n(F_n - 2) + 2 = f(F_n), \end{aligned}$$

where $f(x) = x^2 - 2x + 2$. Hence if $p|F_n$ then $F_{n+1} = f(F_n) \equiv f(0) = 2 \pmod p$, and $F_{n+2} = f(F_{n+1}) \equiv f(2) = 2 \pmod p$; continuing like this we have

$$\begin{aligned} F_{n+k} &= f(f(f(\ldots f(F_{n+1}) \ldots))) \\ &\equiv f(f(f(\ldots f(2) \ldots))) \equiv 2 \pmod p, \end{aligned}$$

[2] It is an open question as to whether there are infinitely many Fermat primes, $F_n$. We have listed the only $F_n$ known to be prime, and for $5 \leq n \leq 30$ the $F_n$ are composite, and for many other $n$ besides. It could be that all $F_n, n > 4$ are composite, or they might all be prime from some sufficiently large $n$ onwards. We have no way of knowing what exactly is true.

for all $k \geq 1$, and so we deduce that $F_m \equiv 2 \pmod{p}$ for all $m > n$. But since each prime factor $p$ of $F_n$ is odd (as $F_n$ is odd), we deduce that the $F_n$ are pairwise coprime. □

When you see two proofs like these last two proofs, that are so similar, you begin to suspect that there may be some deeper unifying idea lying not far below the surface. We explore an appropriate generalization in the next section.

# 2 The arithmetic of dynamical systems

## 2.1 Orbits, periods and pre-periods

We have shown above that the $(a_n)_{n\geq 0}$ and the $(F_n)_{n\geq 0}$ are both examples of sequences $(x_n)_{n\geq 0}$ for which $x_0$ is given and then

$$x_{n+1} = f(x_n) \quad \text{for all } n \geq 0,$$

for some polynomial $f(x) \in \mathbb{Z}[x]$; the $a_n$ with the polynomial $x^2 - x + 1$, and the $F_n$ with the polynomial $x^2 - 2x + 2$. Such sequences are examples of *dynamical systems*, in which the next value of a function depends on its current value. The numbers $(x_n)_{n\geq 0}$ are the *orbit* of $x_0$ under the map $x \to f(x)$. Both proofs used a *period* which means that in, say, the orbit of $y_0$, we have $y_n = y_0$. This implies that $y_{n+j} = y_j$ for all $j \geq 0$, which follows from induction by noting that $y_{n+j+1} = f(y_{n+j}) = f(y_j) = y_{j+1}$.

In our two examples, the key to proving coprimality is that $0$ is *pre-periodic* (i.e. the orbit of $0$ eventually becomes periodic but $0$ is not in the period): For $f(x) = x^2 - x + 1$ we have

$$\boxed{0 \to 1 \to 1 \to \dots},$$

and for $f(x) = x^2 - 2x + 2$ we have

$$\boxed{0 \to 2 \to 2 \to \dots},$$

One can classify the polynomials for which the orbit of $0$ is eventually periodic, and so come up with many more proofs that there are infinitely many primes! There is a big surprise; any period in a dynamical system $x \to f(x)$ with $f(x) \in \mathbb{Z}[x]$ has period length 1 or 2. This can be used to prove that if $0$ is pre-periodic for the map $x \to f(x) \in \mathbb{Z}[x]$ then the orbit of $0$ must be one of the following four basic possibilities (each given here with examples of polynomials for which $0$ has that orbit):

- The polynomial $f(x) = x^2 - ax + a$, indeed any polynomial of the form $a + x(x-a)g(x)$ where $g(x) \in \mathbb{Z}[x]$, has the orbit

$$\boxed{0 \to a \to a \to \dots}$$

- The polynomial $f(x) = x^2 - 2$ gives the case $a = 2$ in the orbit

$$\boxed{0 \to -a \to a \to a \to \dots}$$

One can find such orbits with $a = -2, -1, 1$ or $2$.

- The polynomial $f(x) = x^2 - ax - 1$ gives the case with the minus sign in the orbit

$$\boxed{0 \to \pm 1 \to a \to \pm 1 \to \dots}$$

- The polynomial $f(x) = 1 + x + x^2 - x^3$ gives the case with the plus sign in the orbit

$$\boxed{0 \to \pm 1 \to \pm 2 \to \mp 1 \to \pm 2 \to \dots}$$

(In the last two possibilities one can obtain the case with the other sign by replacing $f(x)$ with $-f(-x)$.)

## 2.2 Proof that all periods have length $1$ or $2$.

Suppose that $N$ is the smallest positive integer for which $a_N = a_0$, so that $a_{N+j} = a_j$ for all $j \geq 0$ (as noted above).

Assume that $N > 1$ so that $a_1 \neq a_0$. Now $a_{n+1} - a_n$ divides $f(a_{n+1}) - f(a_n) = a_{n+2} - a_{n+1}$ for all $n \geq 0$, and so

$a_1 - a_0$ divides $a_2 - a_1$, which divides $a_3 - a_2, \dots,$ which divides $a_N - a_{N-1} = a_0 - a_{N-1}$; and this divides $a_1 - a_N = a_1 - a_0$,

the non-zero number we started with. We deduce that $|a_{j+1} - a_j| = |a_1 - a_0|$ for all $j$. The integers $a_{j+1} - a_j$ cannot all be equal else

$$0 = a_N - a_0 = \sum_{j=0}^{N-1} (a_{j+1} - a_j)$$
$$= \sum_{j=0}^{N-1} (a_1 - a_0) = N(a_1 - a_0) \neq 0,$$

a contradiction. Therefore there must be some $j \geq 1$ for which $a_{j+1} - a_j = -(a_j - a_{j-1})$, and so $a_{j+1} = a_{j-1}$. Therefore $N$, the period length, equals $2$. □

# 3 Final remarks

There are other proofs, many other proofs, that there are infinitely many primes. Some are quite similar to those mentioned here, others are rather different. Some lead to deeper, rich veins of mathematical thought, others are isolated gems, though some are little more than a reformulation of the ideas already known. But it is always a treat to see a new proof and to think through how it fits into the literature and where it leads. Proofs can be found by people at different levels of their education, for example [3] (discussed in section 1.6) was discovered by a student, and is the most original and interesting new proof in years.

Other sources for different proofs of the infinitude of primes include the very popular [2], my own personal favourite, [6], which is a rich (though slightly out-of-date) resource for many things about primes, and the website

`http://www.cut-the-knot.org/proofs/primes.shtml`

Once one knows that there are infinitely many primes, one cannot help but wonder how many are there up to a given point? For example, does the count grow as fast as the count of the number of squares? Or, one might want a big prime and so ask how does one go about finding and identifying primes, and how long should one expect to take to do so?

One might ask whether there are infinitely many primes in a given arithmetic progression; and since the arithmetic progression $a \pmod q$ can be viewed as the values of the polynomial $a + nq$ as $n$ runs through the integers, one might ask whether there are infinitely many prime values of, say, the polynomial $n^2 + 1$, or any other irreducible polynomial.

One might ask whether there is a formula for primes and, if so, is it is a useful formula?

We have answers to some of these questions but not all. And even the answers beg further questions, so that the possibilities are limitless, and always so intriguing.

## References

[1] Takashi Agoh, Paul Erdős, and Andrew Granville, Primes at a (somewhat lengthy) glance. *Amer. Math. Monthly* **104** (1997), 943-945.

[2] Martin Aigner and Günter M. Ziegler, Proofs from The Book (5th edn) Springer-Verlag, Berlin, 2014.

[3] Levent Alpoge, van der Waerden and the primes, *Amer. Math. Monthly* **122** (2015), 784–785.

[4] Enrico Bombieri, Andrew Granville, and Janos Pintz, Squares in arithmetic progressions, *Duke Math. J.* **66** (1992), 369–385.

[5] Harry Furstenberg, On the infinitude of primes. *Amer. Math. Monthly* **62** (1955), 353.

[6] Paulo Ribenboim, The new book of prime number records. Springer-Verlag, New York, 1996.

[7] Brian Rice, Primitive prime divisors in polynomial arithmetic dynamics, *Integers (electronic),* **7:A26** (2007), 16 pages.

### Andrew Granville

Andrew Granville specializes in understanding the distribution of primes, and is co-developer of the *pretentious approach* to analytic number theory. He is co-author of the soon-to-appear graphic novel, *MSI; Anatomy and Permutations* (Princeton University Press, 2018). He is chair of pure mathematics at University College London, as well as the Canadian Research Chair in number theory at the Université de Montréal. This article was developed from the author's 2016 London Taught Course Centre Christmas lecture.