EXAMPLES OF PROOFS BY INDUCTION

KEITH CONRAD

1. Introduction

In this handout we illustrate proofs by induction from several areas of mathematics: linear algebra, polynomial algebra, and calculus. Becoming comfortable with induction proofs is almost entirely a matter of having lots of experience.

The first kind of induction argument one usually sees is for summation identities, such as $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$. This can be proved by induction on n. Familiarity with induction proofs of such identities is *not* adequate preparation for many uses of induction in mathematics.

The reason summation identities are bad models for how induction is used in most of mathematics is that one can't get very far by always trying to "add something to both sides" or similarly manipulate the previously settled cases *into* the next case. Rather than think about the inductive step as deriving the desired result immediately from the inductive hypothesis and algebraic manipulations, it is better to take a broader view: we use the inductive hypothesis (that is, the previously settled cases) to help us derive the next case by some method, which may or may *not* actually start out with the inductive hypothesis as the first part of the argument.

In the proofs below, pay close attention to the point in the inductive step where the inductive hypothesis is actually used.

2. Linear Algebra

Theorem 2.1. *Suppose $B = MAM^{-1}$, where A and B are $n \times n$ matrices and M is an invertible $n \times n$ matrix. Then $B^k = MA^kM^{-1}$ for all integers $k \geq 0$. If A and B are invertible, this equation is true for all integers k.*

Proof. We argue by induction on k, the exponent. (Not on n, the dimension of the matrix!)

The result is obvious for $k = 0$, when both sides equal the $n \times n$ identity matrix I. When $k = 1$, the equation $MA^1M^{-1} = B^1$ is just the original condition $MAM^{-1} = B$.

Let's see what happens in the case when $k = 2$:

$$\begin{aligned} B^2 &= B \cdot B \\ &= (MAM^{-1}) \cdot (MAM^{-1}) \\ &= MA(M^{-1}M)AM^{-1} \\ &= MAIAM^{-1} \\ &= MAAM^{-1} \\ &= MA^2M^{-1} \end{aligned}$$

Now assume the result is established for exponent k. Then

$$\begin{aligned}
B^{k+1} &= B^k \cdot B \\
&= (MA^kM^{-1}) \cdot (MAM^{-1}) \quad \text{(by ind. hyp.)} \\
&= MA^k(M^{-1}M)AM^{-1} \\
&= MA^kIAM^{-1} \\
&= MA^k \cdot AM^{-1} \\
&= MA^{k+1}M^{-1}
\end{aligned}$$

Thus, the result is true for exponent $k+1$ if it is true for exponent k.

Since the base case $k = 1$ is true, and assuming the k-th case is true we showed the $(k+1)$-th case is true, we conclude that the theorem is true for all integers $k \geq 0$.

If A and B are invertible, the result holds for negative exponents as well, since for $k > 0$

$$\begin{aligned}
(MA^kM^{-1}) \cdot (MA^{-k}M^{-1}) &= MA^k(M^{-1}M)A^{-k}M^{-1} \\
&= MA^kIA^{-k}M^{-1} \\
&= MA^kA^{-k}M^{-1} \\
&= MIM^{-1} \\
&= MM^{-1} \\
&= I,
\end{aligned}$$

and thus

$$\begin{aligned}
MA^{-k}M^{-1} &= \text{inverse of } MA^kM^{-1} \\
&= \text{inverse of } B^k \qquad \text{(by the result for } k > 0) \\
&= B^{-k}.
\end{aligned}$$

□

Theorem 2.2. *Let A be a square matrix. Eigenvectors for A having distinct eigenvalues are linearly independent: if $\mathbf{v}_1, \dots, \mathbf{v}_r$ are eigenvectors for A, with $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$ for distinct scalars $\lambda_1, \dots, \lambda_r$, then $\mathbf{v}_1, \dots, \mathbf{v}_r$ are linearly independent.*

Proof. We induct on r, the number of eigenvectors. The result is obvious for $r = 1$. Suppose $r > 1$ and the result has been verified for fewer than r eigenvectors (with distinct eigenvalues).

Given r eigenvectors $\mathbf{v}_i$ of A with distinct eigenvalues λ_i, suppose

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_r\mathbf{v}_r = \mathbf{0} \tag{2.1}$$

for some scalars c_i. We want to prove each c_i is 0.

Applying A to both sides of (2.1), we get

$$c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \cdots + c_r\lambda_r\mathbf{v}_r = \mathbf{0}. \tag{2.2}$$

Now multiply through (2.1) by λ_1:

$$c_1\lambda_1\mathbf{v}_1 + c_2\lambda_1\mathbf{v}_2 + \cdots + c_r\lambda_1\mathbf{v}_r = \mathbf{0}. \tag{2.3}$$

Subtracting (2.3) from (2.2) eliminates the first term:

$$c_2(\lambda_2 - \lambda_1)\mathbf{v}_2 + \cdots + c_r(\lambda_r - \lambda_1)\mathbf{v}_r = \mathbf{0}. \tag{2.4}$$

By the inductive hypothesis, the $r-1$ eigenvectors $\mathbf{v}_2, \dots, \mathbf{v}_r$ are linearly independent. Therefore (2.4) tells us that $c_i(\lambda_i - \lambda_1) = 0$ for $i = 2, 3, \dots, r$. Since the eigenvalues are distinct, $\lambda_i - \lambda_1 \neq 0$, so $c_i = 0$ for $i = 2, 3, \dots, r$. Feeding this into (2.1) gives us $c_1\mathbf{v}_1 = \mathbf{0}$, so $c_1 = 0$ as well. Thus every c_i is 0. □

3. Polynomial algebra

Theorem 3.1. *Let $f(x)$ be a nonconstant polynomial with real coefficients and degree d. Then $f(x)$ has at most d real roots.*

We can't replace "at most d real roots" with "exactly d real roots" since there are nonconstant real polynomials like x^2+1 which have no real roots.

Proof. We induct on the degree d of $f(x)$. Note $d \geq 1$.

A polynomial of degree 1 with real coefficients is of the form $f(x) = ax+b$, where a and b are real and $a \neq 0$. This has exactly one root, namely $-b/a$, and thus *at most* one real root. That settles the theorem for $d = 1$.

Now assume the theorem is true for all polynomials of degree d with real coeffcients. We verify the theorem for all polynomials of degree $d+1$ with real coefficients.

A typical polynomial of degree $d+1$ with real coefficients is

$$f(x) = c_{d+1}x^{d+1} + c_dx^d + \cdots + c_1x + c_0, \tag{3.1}$$

where $c_j \in \mathbf{R}$ and $c_{d+1} \neq 0$. If $f(x)$ has no real roots, then we're done, since $0 \leq d+1$. If $f(x)$ has a real root, say r, then

$$0 = c_{d+1}r^{d+1} + c_dr^d + \cdots + c_1r + c_0. \tag{3.2}$$

Subtracting (3.2) from (3.1), the terms c_0 cancel and we get

$$f(x) = c_{d+1}(x^{d+1} - r^{d+1}) + c_d(x^d - r^d) + \cdots + c_1(x-r) \tag{3.3}$$

Each difference $x^j - r^j$ can has $x - r$ as a factor:

$$x^j - r^j = (x-r)(x^{j-1} + rx^{j-2} + \cdots + r^ix^{j-1-i} + \cdots + r^{j-2}x + r^{j-1}).$$

Write the more complicated second factor, a polynomial of degree $j-1$, as $Q_{j,r}(x)$. So

$$x^j - r^j = (x-r)Q_{j,r}(x), \tag{3.4}$$

and substituting (3.4) into (3.3) gives

$$\begin{aligned} f(x) &= \sum_{j=1}^{d+1} c_j(x-r)Q_{j,r}(x) \\ &= (x-r)\sum_{j=1}^{d+1} c_jQ_{j,r}(x) \\ &= (x-r)(\underbrace{c_{d+1}}_{\neq 0}\overbrace{Q_{d+1,r}(x)}^{\text{degree } d} + \underbrace{\cdots + c_1Q_{1,r}(x)}_{\text{lower degree}}). \end{aligned} \tag{3.5}$$

Since $c_{d+1}Q_{d+1,r}(x)$ is a polynomial of degree d, and each lower degree polynomial does not decrease the degree when added, the second factor in (3.5) has degree d.

Each $Q_{j,r}(x)$ has real coefficients and all c_j are real, so the second factor in (3.5) has real coefficients. We can therefore apply the inductive hypothesis to the second factor and conclude that the second factor has at most d real roots. Any root of $f(x)$ is either r or a root of the second factor, so $f(x)$ has at most $d+1$ real roots. As $f(x)$ was an arbitrary polynomial of degree $d+1$ with real coefficients, we have shown that the d-th case of the theorem being true implies the $(d+1)$-th case is true. By induction on the degree, the theorem is true for all nonconstant polynomials. □

Our next theorem uses the "strong" form of induction, where we argue from all earlier cases to the next case rather than from one case to the next case.

Theorem 3.2. *Every nonconstant polynomial has an irreducible factor.*

Remember that a nonconstant polynomial is called irreducible when its only polynomial factors are constants and constant multiples of itself. (For example, the polynomial x is irreducible. It has factorizations such as $2(x/2)$ or $5(x/5)$, but these are trivial since one of the factors is a constant.)

Proof. We induct on the degree d of the nonconstant polynomial.

When $d = 1$, the polynomial is linear. Linear polynomials are irreducible, so the case $d = 1$ is settled.

Assuming every nonconstant polynomial with degree $\leq d$ has an irreducible factor, consider a polynomial $f(x)$ with degree $d+1$. If $f(x)$ is irreducible, then $f(x)$ has an irreducible factor (namely itself). If $f(x)$ is not irreducible, then we can factor $f(x)$, say $f(x) = g(x)h(x)$ where $g(x)$ and $h(x)$ are nonconstant, so $1 \leq \deg g(x),\ \deg h(x) < d+1$. By our (strong) inductive assumption, $g(x)$ has an irreducible factor, and this irreducible polynomial will also be a factor of $f(x)$ since $g(x)$ is a factor of $f(x)$. Thus $f(x)$ has an irreducible factor and we are done. □

Theorem 3.3. *Let $f(x)$ be a polynomial with integral coefficients and positive degree. For any $k \geq 1$ there is an integer n such that $f(n)$ has at least k different prime factors.*

The meaning of this theorem is that it's impossible for a polynomial with integral coefficients to have its values all be of the form $\pm 2^a 3^b$ or some other product of a *fixed* set of primes.

Proof. This argument is due to Jorge Miranda. We argue by induction on k.

First, since the equations $f(n) = 1$, and $f(n) = -1$ each have only finitely solutions, there are values $f(n)$ divisible by a prime. This settles the case $k = 1$.

Now suppose $k \geq 2$ and there are primes $p_1, \ldots, p_{k-1}$ and a positive integer m such that $f(m)$ is divisible by $p_1, \ldots, p_{k-1}$. We will find a new prime p_k and a value $f(n)$ divisible by $p_1, \ldots, p_{k-1}, p_k$.

If $f(0) = 0$ then $f(n)$ is divisible by n for all n, so letting p_k be a prime other than $p_1, \ldots, p_{k-1}$ the number $f(p_1 \cdots p_k)$ is divisible by $p_1, \ldots, p_k$.

Now suppose $f(0) \neq 0$. Write $f(x) = c_0 + c_1 x + \cdots + c_n x^n = c_0 + xg(x)$, where $c_0 = f(0)$ and $g(x)$ is a nonzero polynomial. Set $f(0) = \pm p_1^{e_1} \cdots p_{k-1}^{e_{k-1}}$. For any n,

$$f(n) = f(0) + ng(n) = \pm p_1^{e_1} \cdots p_{k-1}^{e_{k-1}} + ng(n).$$

If n is divisible by $p_1^{e_1+1} \cdots p_{k-1}^{e_{k-1}+1}$ then the power of each p_i in $f(n)$ is e_i (why?). Therefore $f(n) = f(0)N$ where N is not divisible by any of $p_1, \ldots, p_{k-1}$. The equation $f(n) = \pm f(0)$ has only finitely many solutions, while there are infinitely many multiples of $p_1^{e_1+1} \cdots p_{k-1}^{e_{k-1}+1}$, so there is an n that's a multiple of $p_1^{e_1+1}\cdots p_{k-1}^{e_{k-1}+1}$ such that $f(n)\neq \pm f(0)$. Therefore $f(n)=f(0)N$ where $|N|\geq 2$. A prime factor of N is not any of $p_1,\dots,p_{k-1}$, so $f(n)$ has k prime factors. □

4. Calculus

A calculus analogue of proving summation identities by induction is proving derivative identities by induction. Here is an example.

Theorem 4.1. *For $n\geq 1$, $\dfrac{\mathrm{d}^n}{\mathrm{d}x^n}(e^{x^2})=P_n(x)e^{x^2}$, where $P_n(x)$ is a polynomial of degree n.*

Before working out the proof, let's see why it is easy to *guess* such a result by doing calculations for small n. The first derivative of e^{x^2} is $2xe^{x^2}$. The second derivative of e^{x^2} is $(2xe^{x^2})'$, which is $(4x^2+2)e^{x^2}$ by the product rule. The third derivative of e^{x^2} is $((4x^2+2)e^{x^2})'$, which is $(8x^3+12x)e^{x^2}$ by the product rule again. Each time we are getting a polynomial times e^{x^2}, and the degree of the polynomial matches the order of the derivative. So the formulation of Theorem 4.1 is not a surprise based on these examples.

Proof. We argue by induction on n.

Our base case is $n=0$. The zeroth derivative of a function is the function itself, so we want to know e^{x^2} is $P_0(x)e^{x^2}$ for a polynomial $P_0(x)$ of degree 0. This is true using $P_0(x)=1$. Let's check $n=1$. The first derivative $(e^{x^2})'$ is $2xe^{x^2}$, so this is $P_1(x)e^{x^2}$ for the polynomial $P_1(x)=2x$, which has degree 1.

Now we do the inductive step. For $n\geq 1$, assume

$$(e^{x^2})^{(n)}=P_n(x)e^{x^2} \tag{4.1}$$

for some polynomial $P_n(x)$ of degree n. To compute $(e^{x^2})^{(n+1)}$, we differentiate both sides of (4.1). We obtain

$$\begin{aligned}(e^{x^2})^{(n+1)} &= (P_n(x)e^{x^2})' && \text{(by ind. hyp.)}\\ &= P_n(x)(e^{x^2})'+e^{x^2}P_n'(x) && \text{(by product rule)}\\ &= P_n(x)(2xe^{x^2})+P_n'(x)e^{x^2}\\ &= (2xP_n(x)+P_n'(x))e^{x^2}.\end{aligned}$$

The first factor here is $2xP_n(x)+P_n'(x)$. Since $P_n(x)$ has degree n, $2xP_n(x)$ has degree $n+1$ while $P_n'(x)$ has degree $n-1$. When you add polynomials with different degrees, the degree of the sum is the larger of the two degrees (in fact, the whole leading term of the sum is the leading term of the larger degree polynomial). Therefore, setting $P_{n+1}(x):=2xP_n(x)+P_n'(x)$, we have $(e^{x^2})^{(n+1)}=P_{n+1}(x)e^{x^2}$, where $P_{n+1}(x)$ is a polynomial of degree $n+1$. That settles the inductive step and completes the proof. □

This is a bad example of induction, because the way induction gets used is too routine. Most uses of induction in calculus proofs are not a matter of differentiating both sides of an identity.

Theorem 4.2. *For differentiable functions $f_1(x),\dots,f_n(x)$,*

$$\frac{(f_1(x)\cdots f_n(x))'}{f_1(x)\cdots f_n(x)}=\frac{f_1'(x)}{f_1(x)}+\cdots+\frac{f_n'(x)}{f_n(x)}.$$

Proof. We induct on n, the number of functions.

When $n = 1$, it is clear since both sides equal $f_1'(x)/f_1(x)$.

When $n = 2$, the result follows from the product rule

$$(f_1(x) \cdot f_2(x))' = f_1'(x) \cdot f_2(x) + f_1(x) \cdot f_2'(x)$$

by dividing both sides by $f_1(x)f_2(x)$. (Check this!)

Now assume the result for n functions. When $f_1(x), \dots, f_{n+1}(x)$ are differentiable functions, we write their product as

$$f_1(x)f_2(x)\cdots f_{n+1}(x) = (f_1(x)\cdots f_n(x)) \cdot f_{n+1}(x). \tag{4.2}$$

Considering $f_1(x)\cdots f_n(x)$ as a single function, we can view (4.2) as a product of two functions, and thus use the case $n = 2$:

$$\begin{aligned}
\frac{(f_1(x)\cdots f_{n+1}(x))'}{f_1(x)\cdots f_{n+1}(x)} &= \frac{((f_1(x)\cdots f_n(x)) \cdot f_{n+1}(x))'}{(f_1(x)\cdots f_n(x)) \cdot f_{n+1}(x)} \\
&= \frac{(f_1(x)\cdots f_n(x))'}{f_1(x)\cdots f_n(x)} + \frac{f_{n+1}'(x)}{f_{n+1}(x)} && \text{(Case } n = 2\text{)} \\
&= \frac{f_1'(x)}{f_1(x)} + \cdots + \frac{f_n'(x)}{f_n(x)} + \frac{f_{n+1}'(x)}{f_{n+1}(x)} && \text{(by ind. hyp.)}
\end{aligned}$$

and this is what we needed to show for $n + 1$ functions. □

Remark 4.3. From a logical point of view, the proof above has a gap in the case $n = 2$: we appealed to your familiarity with the product rule rather than include a proof of it. If we were to give a complete proof of the theorem, we should give a proof of the product rule, which would involve going back to the rigorous definition of derivatives involving limits.

Theorem 4.4. *For $n \geq 0$, $\int_0^\infty x^n e^{-x}\,\mathrm{d}x = n!$.*

Proof. We argue by induction on n. For $n = 0$,

$$\int_0^\infty e^{-x}\,\mathrm{d}x = -e^{-x}\Big|_0^\infty = 1.$$

For $n \geq 1$, we express $\int_0^\infty x^n e^{-x}\,\mathrm{d}x$ in terms of $\int_0^\infty x^{n-1} e^{-x}\,\mathrm{d}x$ using integration by parts:

$$\begin{aligned}
\int_0^\infty x^n e^{-x}\,\mathrm{d}x &= \int_0^\infty u\,\mathrm{d}v && (u = x^n, \mathrm{d}v = e^{-x}\,\mathrm{d}x) \\
&= uv\Big|_0^\infty - \int_0^\infty v\mathrm{d}u && (\mathrm{d}u = nx^{n-1}, v = -e^{-x}) \\
&= nx^{n-1}e^{-x}\Big|_0^\infty + n\int_0^\infty x^{n-1}e^{-x}\,\mathrm{d}x \\
&= (0 - 0) + n \cdot (n-1)! && \text{(by ind. hyp.)} \\
&= n!.
\end{aligned}$$

□

Theorem 4.5. *For $x > 0$ and integers $n \geq 0$, $e^x > 1 + x + \dfrac{x^2}{2!} + \cdots + \dfrac{x^n}{n!}$.*

This inequality is clear without induction, using the power series expansion for e^x: $e^x = \sum_{k\geq 0} x^k/k!$ for all real x, and when $x > 0$ the terms in the sum are all positive so we can drop all the terms of the series past the nth term and the inequality of Theorem 4.5 drops out. So why prove Theorem 4.5 by induction if we can prove the theorem quickly using power series? Just to illustrate techniques!

Proof. We will prove the inequality by induction on n.

The base case $n = 0$ says: $e^x > 1$ for $x > 0$. This is true since e^x is an increasing function, so $e^x > e^0 = 1$ when x is positive.

Now we make our inductive hypothesis:

$$e^x > 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}. \tag{4.3}$$

for all $x > 0$. We want to derive the same inequality with $n + 1$ in place of n (again for all $x > 0$).

We will actually give two different approaches to the inductive step: the first will use integrals and the second will use derivatives. These approaches are logically independent of each other and can be read in either order.

The Integral Approach: When $f(x) > g(x)$ on an interval $[a, b]$, their integrals over the interval have the same inequality: $\int_a^b f(x)\mathrm{d}x > \int_a^b g(x)\mathrm{d}x$. This is also true if the functions are equal at an endpoint but otherwise satisfy $f(x) > g(x)$. (We have in mind here only continuous functions.)

We are going to apply this idea to the inequality (4.3). Our inductive hypothesis is that (4.3) holds for every $x > 0$, but at $x = 0$ we get equality in (4.3) since both sides become 1. Therefore we can integrate both sides of (4.3) over *any* interval $[0, t]$ where $t > 0$:

$$\int_0^t e^x \mathrm{d}x > \int_0^t \left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}\right) \mathrm{d}x.$$

Evaluating both sides,

$$e^t - 1 > t + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots + \frac{t^{n+1}}{(n+1)!}.$$

Now add 1 to both sides and we have

$$e^t > 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots + \frac{t^{n+1}}{(n+1)!}.$$

This has been derived for any $t > 0$. So we can now simply rename t as x and we have completed the inductive step.

The Derivative Approach: The key idea we will use with derivatives is that a function having a positive derivative on an interval is increasing on this interval. In particular, if $g(x)$ is differentiable for $x \geq 0$ and $g'(x) > 0$ for $x > 0$ then $g(x) > g(0)$ for $x > 0$. Make sure you understand this idea before reading further.

We are assuming (4.3) holds for some n and want to use this to derive the analogue of (4.3) for the next exponent $n + 1$:

$$e^x \overset{?}{>} 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{n+1}}{(n+1)!}. \tag{4.4}$$

Well, let $F(x)$ be the difference of the two sides of (4.4):

$$F(x) = e^x - \left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!}\right).$$

Our goal is to show $F(x) > 0$ if $x > 0$. Consider the derivative

$$F'(x) = e^x - \left(0 + 1 + x + \cdots + \frac{x^{n-1}}{(n-1)!} + \frac{x^n}{n!}\right).$$

Our induction hypothesis (4.3) is *exactly* the statement that $F'(x)$ is positive for $x > 0$. Therefore by our induction hypothesis $F(x)$ is an increasing function for $x \geq 0$, so $F(x) > F(0)$ when $x > 0$. Since $F(0) = 0$, we obtain $F(x) > 0$ for all $x > 0$, which completes the inductive step using this second approach. $\square$

Notice how the inductive hypothesis was used in the two approaches to the inductive step. In the integral approach, we integrated the inequality in the inductive hypothesis to derive the inequality for the next exponent. In the derivative approach, on the other hand, we did not start with the inductive hypothesis and "do something to both sides." Instead, we set up a convenient function $F(x)$ related to what we wanted to show and used the inductive hypothesis to tell us something relevant about the derivative of that function.

The following corollary of Theorem 4.5 is important: it says e^x grows *faster* than any fixed integral power of x.

Corollary 4.6. *For any integer* $n \geq 0$, $\dfrac{e^x}{x^n} \to \infty$ *as* $x \to \infty$.

Proof. From Theorem 4.5,

$$e^x > 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!}$$

when $x > 0$. (Why did we use the inequality out to degree $n+1$ instead of degree n? Read on and you'll see.) In particular, since all lower degree terms on the right side are positive when $x > 0$,

$$e^x > \frac{x^{n+1}}{(n+1)!}$$

when $x > 0$. Divide both sides of this inequality by x^n:

$$\frac{e^x}{x^n} > \frac{x}{(n+1)!}.$$

Here n is fixed and x is any positive real number. In this inequality, the right side tends to ∞ as $x \to \infty$. Therefore $e^x/x^n \to \infty$ as $x \to \infty$. $\square$

From Corollary 4.6 we draw two further conclusions.

Corollary 4.7. *For any polynomial* $p(x)$, $\dfrac{p(x)}{e^x} \to 0$ *as* $x \to \infty$.

Proof. By Corollary 4.6, for any $n \geq 0$ we have $e^x/x^n \to \infty$ as $x \to \infty$. Inverting the fraction, $x^n/e^x \to 0$ as $x \to \infty$. Any polynomial is a sum of multiples of such ratios: writing $p(x) = a_d x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0$, we have

$$\frac{p(x)}{e^x} = a_d \frac{x^d}{e^x} + a_{d-1}\frac{x^{d-1}}{e^x} + \cdots + a_1 \frac{x}{e^x} + a_0 \frac{1}{e^x}.$$

Each x^n/e^x appearing here tends to 0 as $x \to \infty$, so $p(x)/e^x$ tends to 0 as $x \to \infty$. $\square$

Corollary 4.8. *For any integer $n \geq 0$, $\dfrac{(\log x)^n}{x} \to 0$ as $x \to \infty$,*

Proof. Write $x = e^y$. Then $(\log x)^n/x = y^n/e^y$. We want to show $y^n/e^y \to 0$ as $x \to \infty$. As $x \to \infty$, also $y \to \infty$. Therefore by Corollary 4.6, $e^y/y^n \to \infty$ as $x \to \infty$, so $y^n/e^y \to 0$ as $x \to \infty$. $\square$

Let's return to using induction to prove (interesting) things.

Theorem 4.9. *The exponential function is not "algebraic": there is no $n \geq 1$ and set of polynomials $c_0(x), c_1(x), \ldots, c_n(x)$ not all identically zero such that*

$$c_n(x)e^{nx} + c_{n-1}(x)e^{(n-1)x} + \cdots + c_1(x)e^x + c_0(x) = 0 \tag{4.5}$$

for all $x \in \mathbf{R}$. In other words, if such a functional identity does hold then all the polynomial coefficients $c_k(x)$ are the zero polynomial.

The expression involving e^x in the theorem is a "polynomial with polynomial coefficients," which can be thought of as

$$c_n(x)y^n + c_{n-1}(x)y^{n-1} + \cdots + c_1(x)y + c_0(x) \tag{4.6}$$

where we have substituted e^x for y. Since the $c_k(x)$'s are polynomials in x, (4.6) is a two-variable polynomial in x and y. The theorem is saying no two-variable polynomial $P(x, y)$ can have e^x as a "root" in y. By comparison, the function $x^{1/3}$ is a "root" of the two-variable polynomial $Q(x, y) = y^3 - x$: when we substitute $x^{1/3}$ in for y, the result $Q(x, x^{1/3})$ is the zero function.

Proof. We argue by induction on n. Corollary 4.7 will play a role!

The base case is $n = 1$. Suppose

$$c_1(x)e^x + c_0(x) = 0 \tag{4.7}$$

for all x. We want to show $c_0(x)$ and $c_1(x)$ are both the zero polynomial. Dividing by e^x and re-arranging, we have

$$c_1(x) = -\frac{c_0(x)}{e^x}$$

for all x. We now think about what this tells us as $x \to \infty$. The right side tends to 0 by Corollary 4.7. This forces $c_1(x)$ to be the zero polynomial, since a non-zero polynomial has different limiting behavior as $x \to \infty$: a non-zero constant polynomial keeps its constant value while a non-constant polynomial tends to $\pm\infty$ (depending on the sign of the leading coefficient). Now that we know $c_1(x)$ is identically zero, our original equation (4.7) becomes $c_0(x) = 0$ for all x. This concludes the base case.

For our inductive step, we assume for some n that the only way to satisfy (4.5) for all x is to have all coefficients $c_k(x)$ equal to the zero polynomial. Suppose now that there are polynomials $a_0(x), \ldots, a_{n+1}(x)$ such that

$$a_{n+1}(x)e^{(n+1)x} + a_n(x)e^{nx} + \cdots + a_1(x)e^x + a_0(x) = 0 \tag{4.8}$$

for all x. We want to show every $a_k(x)$ is the zero polynomial.

As in the base case, we divide this equation by an exponential, but now take it to be $e^{(n+1)x}$ instead of e^x:

$$a_{n+1}(x) + \frac{a_n(x)}{e^x} + \cdots + \frac{a_1(x)}{e^{nx}} + \frac{a_0(x)}{e^{(n+1)x}} = 0.$$

Moving all but the first term to the other side,

$$a_{n+1}(x) = -\frac{a_n(x)}{e^x} - \cdots - \frac{a_1(x)}{e^{nx}} - \frac{a_0(x)}{e^{(n+1)x}} \tag{4.9}$$

for all x.

What happens in (4.9) when $x \to \infty$? On the right, $-a_n(x)/e^x \to 0$ by Corollary 4.7. Other terms have the form $-a_k(x)/e^{(n+1-k)x}$ for $k = 0, 1, \ldots, n-1$. Writing this as $-(a_k(x)/e^x)(1/e^{(n-k)x})$, it tends to $0 \cdot 0 = 0$ by Corollary 4.7 (since $n - k > 0$). Therefore the right side of (4.9) tends to 0 as $x \to \infty$, so the polynomial $a_{n+1}(x)$ must be the zero polynomial (same argument as in the base case). This means the left-most term in (4.8) disappears, so (4.8) becomes

$$a_n(x)e^{nx} + \cdots + a_1(x)e^x + a_0(x) = 0$$

for all x. This is a relation of degree n in e^x, so by the inductive hypothesis (at last!) all of its polynomial coefficients are the zero polynomial. Therefore all the coefficients of (4.8) are the zero polynomial, which completes the inductive step. □

Remark 4.10. Since the values of x which mattered in the limits are large values (we took $x \to \infty$), we could have incorporated this into the statement of the theorem and obtained a (slightly) stronger result: if there are polynomials $c_0(x), \ldots, c_n(x)$ such that

$$c_n(x)e^{nx} + c_{n-1}(x)e^{(n-1)x} + \cdots + c_1(x)e^x + c_0(x) = 0$$

just for sufficiently large x, then every $c_k(x)$ is the zero polynomial. The argument proceeds exactly as before except we replace "for all x" by "for all sufficiently large x" in each occurrence. The logical structure of the argument is otherwise unchanged.