*CHAPTER 21 e IS TRANSCENDENTAL

The irrationality of e was so easy to prove that in this optional chapter we will attempt a more difficult feat, and prove that the number e is not merely irrational, but actually much worse. Just how a number might be even worse than irrational is suggested by a slight rewording of definitions. A number x is irrational if it is not possible to write $x = a/b$ for any integers a and b, with $b \neq 0$. This is the same as saying that x does not satisfy any equation

$$bx - a = 0$$

for integers a and b, except for $a = 0$, $b = 0$. Viewed in this light, the irrationality of $\sqrt{2}$ does not seem to be such a terrible deficiency; rather, it appears that $\sqrt{2}$ just barely manages to be irrational—although $\sqrt{2}$ is not the solution of an equation

$$a_1x + a_0 = 0,$$

it *is* the solution of the equation

$$x^2 - 2 = 0,$$

of one higher degree. Problem 2-18 shows how to produce many irrational numbers x which satisfy higher-degree equations

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0,$$

where the a_i are integers not all 0. A number which satisfies an "algebraic" equation of this sort is called an **algebraic number**, and practically every number we have ever encountered is defined in terms of solutions of algebraic equations (π and e are the great exceptions in our limited mathematical experience). All roots, such as

$$\sqrt{2}, \quad \sqrt[10]{3}, \quad \sqrt[4]{7},$$

are clearly algebraic numbers, and even complicated combinations, like

$$\sqrt[3]{3 + \sqrt{5} + \sqrt[4]{1 + \sqrt{2}}} + \sqrt[5]{6}$$

are algebraic (although we will not try to prove this). Numbers which cannot be obtained by the process of solving algebraic equations are called **transcendental**; the main result of this chapter states that e is a number of this anomalous sort.

The proof that e is transcendental is well within our grasp, and was theoretically possible even before Chapter 20. Nevertheless, with the inclusion of this proof, we can justifiably classify ourselves as something more than novices in the study of higher mathematics; while many irrationality proofs depend only on elementary properties of numbers, the proof that a number is transcendental usually involves some really high-powered mathematics. Even the dates connected with the transcendence of e are impressively recent—the first proof that e is transcendental, due to Hermite, dates from 1873. The proof that we will give is a simplification, due to Hilbert.

Before tackling the proof itself, it is a good idea to map out the strategy, which depends on an idea used even in the proof that e is irrational. Two features of the expression

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + R_n$$

were important for the proof that e is irrational: On the one hand, the number

$$1 + \frac{1}{1!} + \cdots + \frac{1}{n!}$$

can be written as a fraction p/q with $q \le n!$ (so that $n!(p/q)$ is an integer); on the other hand, $0 < R_n < 3/(n+1)!$ (so $n!R_n$ is not an integer). These two facts show that e can be approximated particularly well by rational numbers. Of course, every number x can be approximated arbitrarily closely by rational numbers—if $\varepsilon > 0$ there is a rational number r with $|x - r| < \varepsilon$; the catch, however, is that it may be necessary to allow a very large denominator for r, as large as $1/\varepsilon$ perhaps. For e we are assured that this is not the case: there is a fraction p/q within $3/(n+1)!$ of e, whose denominator q is at most $n!$. If you look carefully at the proof that e is irrational, you will see that only this fact about e is ever used. The number e is by no means unique in this respect: generally speaking, the *better* a number can be approximated by rational numbers, the *worse* it is (some evidence for this assertion is presented in Problem 3). The proof that e is transcendental depends on a natural extension of this idea: not only e, but any finite number of powers $e, e^2, \dots, e^n$, can be simultaneously approximated especially well by rational numbers. In our proof we will begin by assuming that e is algebraic, so that

$$(*) \quad a_n e^n + \cdots + a_1 e + a_0 = 0, \qquad a_0 \neq 0$$

for some integers $a_0, \dots, a_n$. In order to reach a contradiction we will than find certain integers $M, M_1, \dots, M_n$ and certain "small" numbers $\epsilon_1, \dots, \epsilon_n$ such that

$$e^1 = \frac{M_1 + \epsilon_1}{M},$$
$$e^2 = \frac{M_2 + \epsilon_2}{M},$$
$$\vdots$$
$$e^n = \frac{M_n + \epsilon_n}{M}.$$

Just how small the ϵ's must be will appear when these expressions are substituted into the assumed equation $(*)$. After multiplying through by M we obtain

$$[a_0 M + a_1 M_1 + \cdots + a_n M_n] + [\epsilon_1 a_1 + \cdots + \epsilon_n a_n] = 0.$$

The first term in brackets is an integer, and we will choose the M's so that it will necessarily be a *nonzero* integer. We will also manage to find ϵ's so small that

$$|\epsilon_1 a_1 + \cdots + \epsilon_n a_n| < \tfrac{1}{2};$$

this will lead to the desired contradiction—the sum of a nonzero integer and a number of absolute value less than $\frac{1}{2}$ cannot be zero!

As a basic strategy this is all very reasonable and quite straightforward. The remarkable part of the proof will be the way that the M's and ϵ's are defined. In order to read the proof you will need to know about the gamma function! (This function was introduced in Problem 19-40.)

THEOREM 1 e is transcendental.

PROOF Suppose there were integers $a_0, \dots, a_n$, with $a_0 \neq 0$, such that

$$(*) \quad a_n e^n + a_{n-1} e^{n-1} + \cdots + a_0 = 0.$$

Define numbers $M, M_1, \dots, M_n$ and $\epsilon_1, \dots, \epsilon_n$ as follows:

$$M = \int_0^\infty \frac{x^{p-1}[(x-1)\cdot\ldots\cdot(x-n)]^p e^{-x}}{(p-1)!}\,dx,$$

$$M_k = e^k \int_k^\infty \frac{x^{p-1}[(x-1)\cdot\ldots\cdot(x-n)]^p e^{-x}}{(p-1)!}\,dx,$$

$$\epsilon_k = e^k \int_0^k \frac{x^{p-1}[(x-1)\cdot\ldots\cdot(x-n)]^p e^{-x}}{(p-1)!}\,dx.$$

The unspecified number p represents a prime number* which we will choose later. Despite the forbidding aspect of these three expressions, with a little work they will appear much more reasonable. We concentrate on M first. If the expression in brackets,

$$[(x-1)\cdot\ldots\cdot(x-n)],$$

is actually multiplied out, we obtain a polynomial

$$x^n + \cdots \pm n!$$

*The term "prime number" was defined in Problem 2-17. An important fact about prime numbers will be used in the proof, although it is not proved in this book: If p is a prime number which does not divide the integer a, and which does not divide the integer b, then p also does not divide ab. The Suggested Reading mentions references for this theorem (which is crucial in proving that the factorization of an integer into primes is unique). We will also use the result of Problem 2-17(d), that there are infinitely many primes—the reader is asked to determine at precisely which points this information is required.

with integer coefficients. When raised to the pth power this becomes an even more complicated polynomial

$$x^{np} + \cdots \pm (n!)^p.$$

Thus M can be written in the form

$$M = \sum_{\alpha=0}^{np} \frac{1}{(p-1)!} C_\alpha \int_0^\infty x^{p-1+\alpha} e^{-x}\, dx,$$

where the C_α are certain integers, and $C_0 = \pm(n!)^p$. But

$$\int_0^\infty x^k e^{-x}\, dx = k!.$$

Thus

$$M = \sum_{\alpha=0}^{np} C_\alpha \frac{(p-1+\alpha)!}{(p-1)!}.$$

Now, for $\alpha = 0$ we obtain the term

$$\pm(n!)^p \frac{(p-1)!}{(p-1)!} = \pm(n!)^p.$$

We will now consider only primes $p > n$; then this term is an integer which is *not* divisible by p. On the other hand, if $\alpha > 0$, then

$$C_\alpha \frac{(p-1+\alpha)!}{(p-1)!} = C_\alpha (p+\alpha-1)(p+\alpha-2) \cdot \ldots \cdot p,$$

which *is* divisible by p. Therefore M itself is an integer which is *not* divisible by p.

Now consider M_k. We have

$$\begin{aligned} M_k &= e^k \int_k^\infty \frac{x^{p-1}[(x-1) \cdot \ldots \cdot (x-n)]^p e^{-x}}{(p-1)!}\, dx \\ &= \int_k^\infty \frac{x^{p-1}[(x-1) \cdot \ldots \cdot (x-n)]^p e^{-(x-k)}}{(p-1)!}\, dx. \end{aligned}$$

This can be transformed into an expression looking very much like M by the substitution

$$\begin{aligned} u &= x - k \\ du &= dx. \end{aligned}$$

The limits of integration are changed to 0 and ∞, and

$$M_k = \int_0^\infty \frac{(u+k)^{p-1}[(u+k-1) \cdot \ldots \cdot u \cdot \ldots \cdot (u+k-n)]^p e^{-u}}{(p-1)!}\, du.$$

There is one very significant difference between this expression and that for M. The term in brackets contains the factor u in the kth place. Thus the pth power contains the factor u^p. This means that the entire expression

$$(u+k)^{p-1}[(u+k-1) \cdot \ldots \cdot (u+k-n)]^p$$

is a polynomial with integer coefficients, *every term of which* has degree at least p. Thus

$$M_k = \sum_{\alpha=1}^{np} \frac{1}{(p-1)!} D_\alpha \int_0^\infty u^{p-1+\alpha} e^{-u}\, du = \sum_{\alpha=1}^{np} D_\alpha \frac{(p-1+\alpha)!}{(p-1)!},$$

where the D_α are certain integers. Notice that the summation begins with $\alpha = 1$; in this case *every* term in the sum is divisible by p. Thus each M_k is an integer which *is* divisible by p.

Now it is clear that

$$e^k = \frac{M_k + \epsilon_k}{M}, \qquad k = 1, \dots, n.$$

Substituting into $(*)$ and multiplying by M we obtain

$$[a_0 M + a_1 M_1 + \cdots + a_n M_n] + [a_1 \epsilon_1 + \cdots + a_n \epsilon_n] = 0.$$

In addition to requiring that $p > n$ let us also stipulate that $p > |a_0|$. This means that both M and a_0 are not divisible by p, so $a_0 M$ is also not divisible by p. Since each M_k is divisible by p, it follows that

$$a_0 M + a_1 M_1 + \cdots + a_n M_n$$

is *not* divisible by p. In particular it is a *nonzero* integer.

In order to obtain a contradiction to the assumed equation $(*)$, and thereby prove that e is transcendental, it is only necessary to show that

$$|a_1 \epsilon_1 + \cdots + a_n \epsilon_n|$$

can be made as small as desired, by choosing p large enough; it is clearly sufficient to show that each $|\epsilon_k|$ can be made as small as desired. This requires nothing more than some simple estimates; for the remainder of the argument remember that n is a certain fixed number (the degree of the assumed polynomial equation $(*)$). To begin with, if $1 \le k \le n$, then

$$\begin{aligned}
|\epsilon_k| &\le e^k \int_0^k \frac{|x^{p-1}[(x-1)\cdot\ldots\cdot(x-n)]^p|\, e^{-x}}{(p-1)!}\, dx \\
&\le e^n \int_0^n \frac{n^{p-1}|(x-1)\cdot\ldots\cdot(x-n)|^p e^{-x}}{(p-1)!}\, dx.
\end{aligned}$$

Now let A be the maximum of $|(x-1)\cdot\ldots\cdot(x-n)|$ for x in $[0, n]$. Then

$$\begin{aligned}
|\epsilon_k| &\le \frac{e^n n^{p-1} A^p}{(p-1)!} \int_0^n e^{-x}\, dx \\
&\le \frac{e^n n^{p-1} A^p}{(p-1)!} \int_0^\infty e^{-x}\, dx \\
&= \frac{e^n n^{p-1} A^p}{(p-1)!} \\
&\le \frac{e^n n^p A^p}{(p-1)!} = \frac{e^n (nA)^p}{(p-1)!}.
\end{aligned}$$

But n and A are fixed; thus $(nA)^p/(p-1)!$ can be made as small as desired by making p sufficiently large. ▮

This proof, like the proof that π is irrational, deserves some philosophic afterthoughts. At first sight, the argument seems quite "advanced"—after all, we use integrals, and integrals from 0 to ∞ at that. Actually, as many mathematicians have observed, integrals can be eliminated from the argument completely; the only integrals essential to the proof are of the form

$$\int_0^\infty x^k e^{-x}\,dx$$

for integral k, and these integrals can be replaced by $k!$ whenever they occur. Thus M, for example, could have been defined initially as

$$M = \sum_{\alpha=0}^{np} C_\alpha \frac{(p-1+\alpha)!}{(p-1)!},$$

where C_α are the coefficients of the polynomial

$$[(x-1)\cdot\ldots\cdot(x-n)]^p.$$

If this idea is developed consistently, one obtains a "completely elementary" proof that e is transcendental, depending only on the fact that

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots.$$

Unfortunately, this "elementary" proof is harder to understand than the original one—the whole structure of the proof must be hidden just to eliminate a few integral signs! This situation is by no means peculiar to this specific theorem—"elementary" arguments are frequently more difficult than "advanced" ones. Our proof that π is irrational is a case in point. You probably remember nothing about this proof except that it involves quite a few complicated functions. There is actually a more advanced, but much more conceptual proof, which shows that π is *transcendental*, a fact which is of great historical, as well as intrinsic, interest. One of the classical problems of Greek mathematics was to construct, with compass and straightedge alone, a square whose area is that of a circle of radius 1. This requires the construction of a line segment whose length is $\sqrt{\pi}$, which can be accomplished if a line segment of length π is constructible. The Greeks were totally unable to decide whether such a line segment could be constructed, and even the full resources of modern mathematics were unable to settle this question until 1882. In that year Lindemann proved that π is transcendental; since the length of any segment that can be constructed with straightedge and compass can be written in terms of $+$, $\cdot$, $-$, $\div$, and $\sqrt{}$, and is therefore algebraic, this proves that a line segment of length π cannot be constructed.

The proof that π is transcendental requires a sizable amount of mathematics which is too advanced to be reached in this book. Nevertheless, the proof is not much more difficult than the proof that e is transcendental. In fact, the proof for π is practically the same as the proof for e. This last statement should certainly surprise you. The proof that e is transcendental seems to depend so thoroughly on particular properties of e that it is almost inconceivable how any modifications could ever be used for π; after all, what does e have to do with π? Just wait and see!

PROBLEMS

1. (a) Prove that if $\alpha > 0$ is algebraic, then $\sqrt{\alpha}$ is algebraic.
(b) Prove that if α is algebraic and r is rational, then $\alpha + r$ and αr are algebraic.

Part (b) can actually be strengthened considerably: the sum, product, and quotient of algebraic numbers is algebraic. This fact is too difficult for us to prove here, but some special cases can be examined:

2. Prove that $\sqrt{2} + \sqrt{3}$ and $\sqrt{2}(1 + \sqrt{3})$ are algebraic, by actually finding algebraic equations which they satisfy. (You will need equations of degree 4.)

***3.** (a) Let α be an algebraic number which is not rational. Suppose that α satisfies the polynomial equation

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0,$$

and that no polynomial function of lower degree has this property. Show that $f(p/q) \neq 0$ for any rational number p/q. Hint: Use Problem 3-7(b).
(b) Now show that $|f(p/q)| \geq 1/q^n$ for all rational numbers p/q with $q > 0$. Hint: Write $f(p/q)$ as a fraction over the common denominator q^n.
(c) Let $M = \sup\{|f'(x)| : |x - \alpha| < 1\}$. Use the Mean Value Theorem to prove that if p/q is a rational number with $|\alpha - p/q| < 1$, then $|\alpha - p/q| > 1/Mq^n$. (It follows that for $c = \min(1, 1/M)$ we have $|\alpha - p/q| > c/q^n$ for all rational p/q.)

***4.** Let

$$\alpha = 0.110001000000000000000001000\ldots,$$

where the 1's occur in the $n!$ place, for each n. Use Problem 3 to prove that α is transcendental. (For each n, show that α is not the root of an equation of degree n.)

Although Problem 4 mentions only one specific transcendental number, it should be clear that one can easily construct infinitely many other numbers α which do not satisfy $|\alpha - p/q| > c/q^n$ for any c and n. Such numbers were first considered by Liouville (1809–1882), and the inequality in Problem 3 is often called Liouville's inequality. None of the transcendental numbers constructed in this way happens to be particularly interesting, but for a long time Liouville's transcendental numbers were the only ones known. This situation was changed quite radically by the work of Cantor (1845–1918), who showed, without exhibiting a single transcendental number, that *most* numbers are transcendental. The next two problems provide an introduction to the ideas that allow us to make sense of such statements. The basic definition with which we must work is the following: A set A is called **countable** if its elements can be arranged in a sequence

$$a_1, a_2, a_3, a_4, \ldots.$$

The obvious example (in fact, more or less the Platonic ideal of) a countable set is $\mathbf{N}$, the set of natural numbers; clearly the set of even natural numbers is also countable:

$$2, 4, 6, 8, \ldots.$$

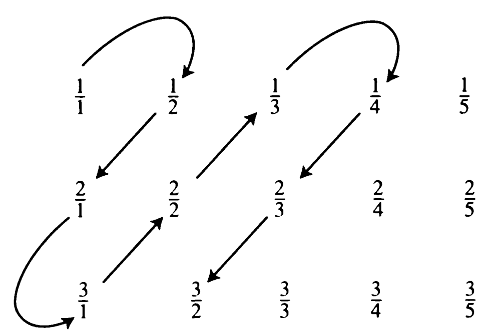
It is a little more surprising to learn that $\mathbf{Z}$, the set of all integers (positive, negative and 0) is also countable, but seeing is believing:

$$0, 1, -1, 2, -2, 3, -3, \ldots.$$

The next two problems, which outline the basic features of countable sets, are really a series of examples to show that (1) a lot more sets are countable than one might think and (2) nevertheless, some sets are not countable.

***5.** (a) Show that if A and B are countable, then so is $A \cup B = \{x : x \text{ is in } A \text{ or } x \text{ is in } B\}$. Hint: Use the same trick that worked for $\mathbf{Z}$.

(b) Show that the set of positive rational numbers is countable. (This is really quite startling, but the figure below indicates the path to enlightenment.)



(c) Show that the set of all pairs (m, n) of integers is countable. (This is practically the same as part (b).)

(d) If $A_1, A_2, A_3, \ldots$ are each countable, prove that

$$A_1 \cup A_2 \cup A_3 \cup \ldots$$

is also countable. (Again use the same trick as in part (b).)

(e) Prove that the set of all triples (l, m, n) of integers is countable. (A triple (l, m, n) can be described by a pair (l, m) and a number n.)

(f) Prove that the set of all n-tuples $(a_1, a_2, \ldots, a_n)$ is countable. (If you have done part (c), you can do this, using induction.)

(g) Prove that the set of all roots of polynomial functions of degree n with integer coefficients is countable. (Part (f) shows that the set of all these polynomial functions can be arranged in a sequence, and each has at most n roots.)

(h) Now use parts (d) and (g) to prove that the set of all algebraic numbers is countable.

***6.** Since so many sets turn out to be countable, it is important to note that the set of all real numbers between 0 and 1 is *not* countable. In other words, there is no way of listing all these real numbers in a sequence

$$\begin{aligned} \alpha_1 &= 0.a_{11}a_{12}a_{13}a_{14}\dots \\ \alpha_2 &= 0.a_{21}a_{22}a_{23}a_{24}\dots \\ \alpha_3 &= 0.a_{31}a_{32}a_{33}a_{34}\dots \\ &\vdots \end{aligned}$$

(decimal notation is being used on the right). To prove that this is so, suppose such a list were possible and consider the decimal

$$0.\bar{a}_{11}\bar{a}_{22}\bar{a}_{33}\bar{a}_{44}\dots,$$

where $\bar{a}_{nn} = 5$ if $a_{nn} \neq 5$ and $\bar{a}_{nn} = 6$ if $a_{nn} = 5$. Show that this number cannot possibly be in the list, thus obtaining a contradiction.

Problems 5 and 6 can be summed up as follows. The set of algebraic numbers is countable. If the set of transcendental numbers were also countable, then the set of all real numbers would be countable, by Problem 5(a), and consequently the set of real numbers between 0 and 1 would be countable. But this is false. Thus, the set of algebraic numbers is countable and the set of transcendental numbers is not ("there are more transcendental numbers than algebraic numbers"). The remaining two problems illustrate further how important it can be to distinguish between sets which are countable and sets which are not.

***7.** Let f be a nondecreasing function on $[0, 1]$. Recall (Problem 8-8) that $\lim_{x\to a^+} f(x)$ and $\lim_{x\to a^-} f(x)$ both exist.

(a) For any $\varepsilon > 0$ prove that there are only finitely many numbers a in $[0, 1]$ with $\lim_{x\to a^+} f(x) - \lim_{x\to a^-} f(x) > \varepsilon$. Hint: There are, in fact, at most $[f(1) - f(0)]/\varepsilon$ of them.

(b) Prove that the set of points at which f is discontinuous is countable. Hint: If $\lim_{x\to a^+} f(x) - \lim_{x\to a^-} f(x) > 0$, then it is $> 1/n$ for some natural number n.

This problem shows that a nondecreasing function is automatically continuous at most points. For differentiability the situation is more difficult to analyze and also more interesting. A nondecreasing function can fail to be differentiable at a set of points which is not countable, but it is still true that nondecreasing functions are differentiable at most points (in a different sense of the word "most"). Reference [38] of the Suggested Reading gives a proof using the Rising Sun Lemma of Problem 8-20.

For those who have done Problem 9 of the Appendix to Chapter 11, it is possible to provide at least one application to differentiability of the ideas already developed in this problem set: If f is convex, then f is differentiable except at those points where its right-hand derivative $f_+{}'$ is discontinuous; but the function $f_+{}'$ is increasing, so a convex function is automatically differentiable except at a countable set of points.

***8.** (a) Problem 11-70 showed that if every point is a local maximum point for a *continuous* function f, then f is a constant function. Suppose now that the hypothesis of continuity is dropped. Prove that f takes on only a countable set of values. Hint: For each x choose *rational* numbers a_x and b_x such that $a_x < x < b_x$ and x is a maximum point for f on (a_x, b_x). Then every value $f(x)$ is the maximum value of f on some interval (a_x, b_x). How many such intervals are there?

(b) Deduce Problem 11-70(a) as a corollary.

(c) Prove the result of Problem 11-70(b) similarly.