Variations on an Irrational Theme —Geometry, Dynamics, Algebra

DAN KALMAN
American University
Washington, DC 20016

ROBERT MENA
California State University
Long Beach, CA 90840

SHAHRIAR SHAHRIARI
Pomona College
Claremont, CA 91711

If someone mentions *irrational number*, what do you think of? Perhaps you recall a standard example, $\sqrt{2}$, and a proof by contradiction that has to do with odd and even numbers. Or perhaps what comes to mind is that the Pythagoreans were discomfitted by the irrationality of $\sqrt{2}$ because it proved that not all geometric relationships could be described in terms of whole numbers. In this paper we will touch on both of these aspects of irrationality, recounting a bit of the history, and showing some variations on the traditional approaches to these topics. Although the subject is a familiar one, it is rich in interesting ideas. The purpose of this paper is to popularize some irrational ideas that do not appear to be well known, including connections to eigenvalues and dynamical systems, and to bring them together with some of the ideas that are so familiar.

Incommensurability and Infinite Descent

The Pythagoreans encountered the idea of irrationality in geometry in the context of commensurability. Initially, in harmony with their *all is number* doctrine, they embraced the geometric position that any two segments are commensurable, meaning, exactly measurable with a common unit. In modern terms, that would mean that relative to an arbitrary unit of measurement, every segment has rational length. Of course that is false, and the very notion seems quaint to our ears. But it was an unexpected discovery to the Greeks, and had fundamental mathematical and philosophical ramifications. According to one oft-repeated account, the demonstration of the existence of incommensurable segments was so devastating that the bearer of the bad news was put to death for his discovery.[1]

To understand the importance of commensurability to the Pythagoreans, one must bear in mind their reliance on whole number relationships. In particular, the concept of proportion was formulated in integral terms: the fundamental observation is that $a:b$ and $na:nb$ are in equal proportion. Then clearly $ma:mb$ and $na:nb$ are also equal. In geometry, with the quantities ma and na representing line segments, the common divisor a becomes a common unit of measurement.

[1] As retold by Choike [4], the discoverer, Hippasus of Metapontum, was on a voyage at the time, and his fellows cast him overboard. A more restrained discussion by Boyer [2, pp. 71–72] describes both the discovery by Hippasus and his execution by drowning as mere possibilities.

Proportionality of Similar Triangles As a concrete example of this idea, we will derive the proportionality of the corresponding parts of similar triangles, following the approach of Aaboe [**1**, pp. 42–43]. Let ABC and $A'B'C'$ be triangles whose corresponding angles are equal, and suppose that BC and $B'C'$ are measured by the common unit a. Then for some integers n and m, $BC = ma$ and $B'C' = na$, as illustrated in FIGURE 1.

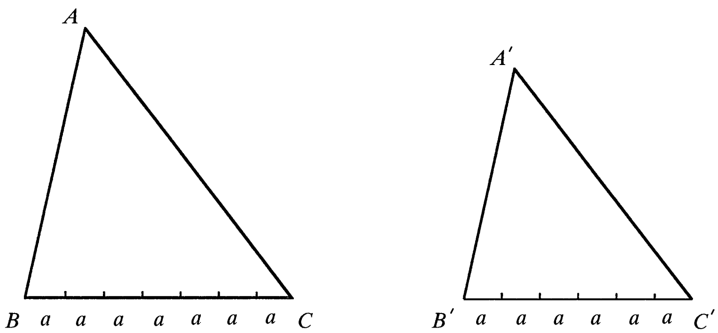

FIGURE 1
Triangles with commensurable bases

Focusing on ABC for a moment, observe that the subdivision of BC into m equal segments permits us also to subdivide AC into m equal segments: simply construct parallel lines as shown in FIGURE 2.

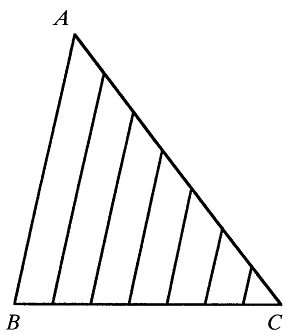

FIGURE 2
Subdividing AC

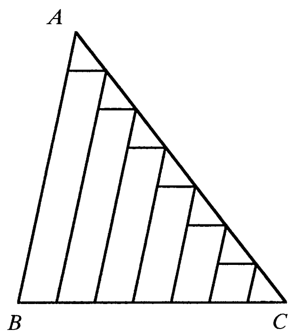

FIGURE 3
Parallelograms

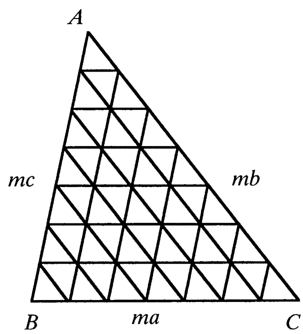

FIGURE 4
ABC Tiled

The intersections of these parallel lines with AC are equally spaced along that side. This can be seen by constructing line segments parallel to BC as in FIGURE 3. Each segment has length a, because it completes a parallelogram with base of length a along BC. That makes the triangles lying along AC congruent, and so verifies that their sides on AC are all of equal length, say b.

And now with two sides of the triangle subdivided, we can partition the remaining side into m equal parts of length c in two ways, using lines parallel to either AC or BC. The result is actually a tiling of ABC by congruent triangles, with m tiles along each side (FIGURE 4). In each tile, the sides are a, b, and c. Thus $AB = mc$ and $AC = mb$.

The same construction carried out in $A'B'C'$ results in a tiling with n copies of the tile along each side (FIGURE 5). Moreover, the tiles used in each triangle are congruent. By construction they clearly share equal corresponding angles, as well as one side, a. This leads to $A'B' = nc$ and $A'C' = nb$, and proves that the sides of the triangles are in equal proportion. For example, $BC : B'C' = ma : na = mb : nb = AC : A'C'$.

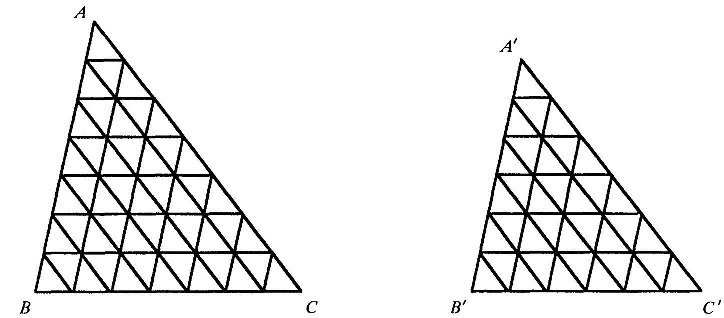

FIGURE 5
Tiled Triangles

If it is assumed that all pairs of segments are commensurable, this argument establishes the proportionality of similar triangles. More generally, the presumption of commensurability justifies treating all proportions as ratios of integers. The discovery of incommensurable segments revealed a fundamental flaw in this approach to proportionality, and led ultimately to the much more sophisticated formulation that appears in Book V of Euclid.

Infinite Descent No one really knows how incommensurability was first discovered. In [4], there is a retelling of the suggestion of von Fritz [9] that the pentagram was the first geometric figure shown to have incommensurable parts. The argument given there uses the idea of infinite descent. Starting with an initial figure, we construct another similar figure that is demonstrably smaller in size. Two parts of the original figure are assumed to be measurable with a common unit, and then it is shown that this same unit must measure the corresponding parts of the smaller figure. By repeating the construction, we can eventually reduce the figure so far that the diameter is less than the common unit, whereupon we contradict the fact that this unit must measure two sides of the figure. In [4] this argument is made using a pentagram. Here we will give a somewhat simpler construction starting with an isosceles right triangle. An essentially equivalent construction, working in a square, is presented in [3].

Consider FIGURE 6, showing an isosceles right triangle ABC. The point D has been constructed so that $BD = BC$. Through D we draw a line parallel to leg AC, which meets BC at point E. Now construct a square having CE as one side (see FIGURE 7), thus defining points F and G.

For reference, we have drawn the auxiliary lines CG and CD in FIGURE 8. Observe now that CG and GD have equal length. Indeed, with BC and BD equal (by the construction of D), we know that angles DCB and CDB are equal. Also angles GCE and GDB are equal (and each is half a right angle). Thus, triangle CDG is isosceles, with CG and GD equal, as asserted. To complete the construction, add point H to define a parallelogram $ADGH$ (FIGURE 9). Then triangles FGH and FGC are congruent, so that CG and GH are equal. Combined with the earlier result, this shows that all sides of parallelogram $ADGH$ are equal to CG.

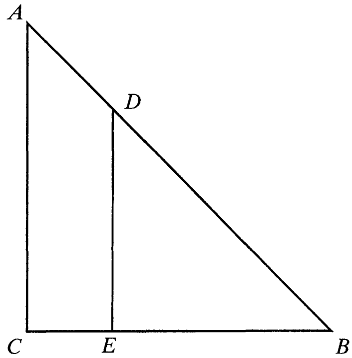

FIGURE 6
Isosceles right triangle

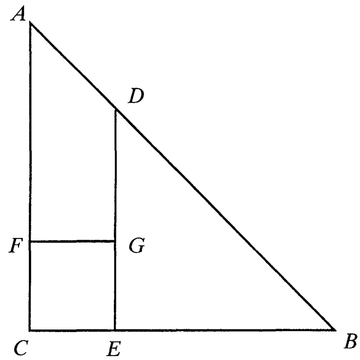

FIGURE 7
CEGF is a square

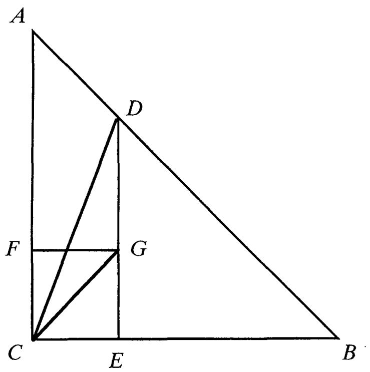

FIGURE 8
Triangles *BCD* and *GCD* are isosceles

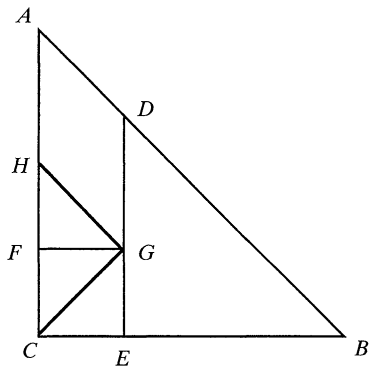

FIGURE 9
ADGH is a parallelogram

To summarize the result of the construction, FIGURE 10 shows the essential segments, with AH, AD, HG, and GC all equal in length. Triangle CGH is an isosceles right triangle. If a unit evenly measures BC and AB, then it must also measure their difference, AD. The unit therefore measures legs HG and CG of CGH. Furthermore, since the unit measures both AC and AH, it measures their difference, CH, the hypotenuse of CGH. Therefore, any unit that measures the parts

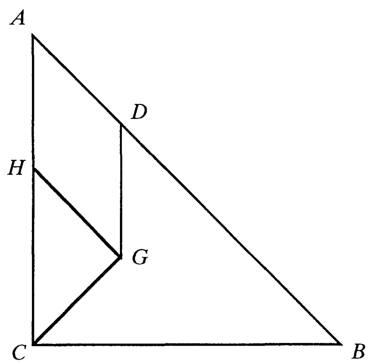

FIGURE 10
A unit measuring AB and BC also measures CG and CH

of triangle ABC must also measure the parts of the smaller similar triangle CGH. This completes the construction. The incommensurability of AB and BC now follows as discussed earlier.

The incommensurability argument also leads to an algebraic demonstration of the irrationality of $\sqrt{2}$. Assume that there is a unit that divides evenly into the leg and the hypotenuse of the original triangle, say with n units along AB and m on BC. Then AD must be measured by $n-m$ units, as must AH, GH, and GC. Furthermore, HC is then measured by $m-(n-m)=2m-n$ units. Since CHG and ABC are similar triangles, we conclude that $n/m=(2m-n)/(n-m)$. This same conclusion can be reached using algebra. Suppose that a and b satisfy $a^2=2b^2$. Then $a^2-ab=2b^2-ab$ hence $a(a-b)=b(2b-a)$. This leads to our earlier conclusion: $a/b=(2b-a)/(a-b)$. Now since $b<a$, we see that $2b-a<a$. Since $a<2b$, $a-b<b$. That is, the numerator and denominator of $(2b-a)/(a-b)$ are each less than the corresponding parts of a/b. The conclusion is summarized as follows: Any ratio a/b representing $\sqrt{2}$ leads to another ratio with strictly smaller numerator and denominator. If a and b are integers, so are $2b-a$ and $a-b$. Thus, given any integer ratio for $\sqrt{2}$ we obtain an equal ratio of strictly smaller integers. This is clearly an impossible situation, so $\sqrt{2}$ must have no such representation.

The preceding argument appears in [**10**, pp. 39–41]. It is essentially the same as one used by Fermat to argue the irrationality of $\sqrt{3}$ (see [**2**, pp. 353–354]). Fermat went on to make great use of the notion of infinite descent in number theory. In contrast, our discourse now heads in a different direction—to the use of matrices.

A Dynamical View of Irrationality

One facet that both the algebraic and geometric infinite descent arguments share is the propagation of pairs (a,b). Indeed, the generation of each new pair from its predecessor is of a linear nature. It is natural therefore to represent it as a matrix operation. Let A be the matrix $\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$, and represent the pair (a,b) as a column vector. Then

$$\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}=\begin{bmatrix} 2b-a \\ a-b \end{bmatrix}$$

describes the propagation used in our earlier arguments. Now we make two observations about A. First, as an integer matrix, it preserves lattice points. That is, if v is a vector with integer components, then so is Av. Second, the line L described by $a=\sqrt{2}\,b$ is an eigenspace, so its points are also preserved by A. Actually we can say more: A is a contraction on L. Simply observe that

$$\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix}=\begin{bmatrix} 2-\sqrt{2} \\ \sqrt{2}-1 \end{bmatrix}=\begin{bmatrix} \sqrt{2}\,(\sqrt{2}-1) \\ \sqrt{2}-1 \end{bmatrix}=(\sqrt{2}-1)\begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix}$$

Since the eigenvalue $\sqrt{2}-1$ is between 0 and 1, the effect of A on points of L is to reduce their magnitude.

The infinite descent argument can now be stated in dynamical terms. Starting with any first quadrant point (a,b) on L, repeated application of A generates a sequence of points that remain on the line while converging to 0. If the initial point were a lattice point, all of the successive points would be as well, leading us to the impossible situation of an infinite sequence of distinct lattice points converging to the origin. We conclude that there are no lattice points on L.

Dynamics of A and A^{-1} There is a bigger dynamical picture. Although there are no lattice points on L, there are plenty elsewhere in the plane. Repeated application of A to each must generate a sequence of lattice points, called an *orbit*. Where do these orbits lead? It is easy to show that A has another eigenvalue with magnitude greater than 1, and a corresponding line M of eigenvectors. Each element of the plane can be expressed as a sum of elements of L and M. Under repeated application of A, the L component dwindles away to nothing, while the M component grows without bound. Therefore, almost all of the points in the plane, including every one of the lattice points, march off to infinity under the action of A. This is the dynamical systems view of A. Its repeated application to the plane sweeps everything not on L out to infinity along M, while the points on L all flow toward the origin. In combination with the fact that A preserves integer lattice points, this shows that L can contain no lattice points other than 0.

Somewhat paradoxically, although the dynamic description is given in very geometric terms, it is not easy to depict accurately on a graph. For one thing, the eigenvalue corresponding to M is negative. As A is repeatedly applied to a vector, the M component alternates in sign. The resulting orbit jumps back and forth, progressing in one direction along M on the even jumps, and in the opposite direction on the odd jumps. So "marching to infinity" is not really the right image. Rather, the points leap-frog infinitely far along M in both directions. Looking just at the landing points of the even leaps, the points seem to follow a flow, as illustrated qualitatively in FIGURE 11. This really shows the dynamic behavior of A^2. It gives some sense of the dynamics of A, as long as you remember what is happening on the odd leaps.

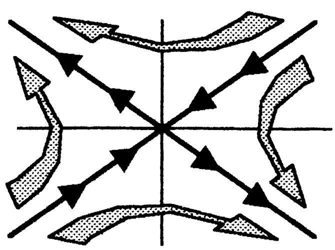

FIGURE 11
Dynamics of A^2

The magnitude of the negative eigenvalue presents another obstacle to forming an accurate graphical representation of the dynamics of A. Except for points very close to L, the M component grows so rapidly that the L component becomes completely invisible after only one or two iterations. That is, if the scale is made large enough to show an initial point and two iterates, relative to that scale, even the initial L component will be hard to see. This effect is illustrated in FIGURE 12, which shows a square, and its images under A and A^2. The second image is hard to distinguish from a heavily inked line. Careful inspection reveals the effects of the negative eigenvalue, as the labeled vertices alternate orientation around the square and its successive images. However, with only two applications of A illustrated, there is not much of a basis for visualizing the overall structure of the orbits. In fact, the situation is more easily described than drawn. From just about any starting point, the orbit takes only a step or two to get right next to M. From that point on, the orbit jumps off to infinity, alternating between one end of M and the other.

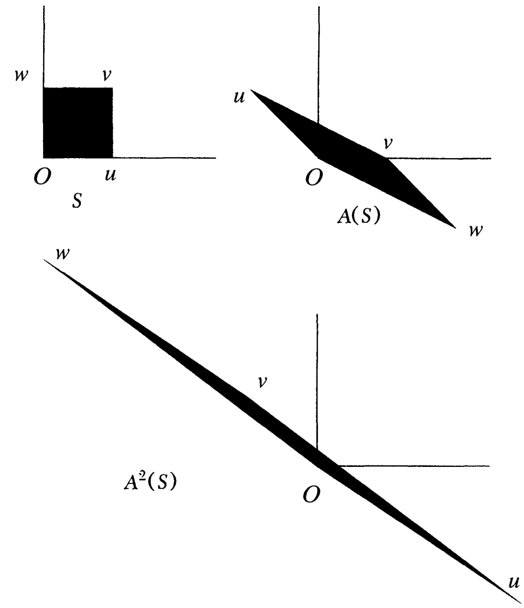

FIGURE 12
Applications of A to a square S

As stated earlier, A carries each lattice point to another lattice point. As a matter of fact, the set of lattice points is actually invariant under A, because the inverse $\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ also has integer entries. The dynamics of A^{-1} are the reverse of those of A: all the points off M are swept out to infinity along L, while points on M collapse into the origin. Reasoning exactly as before, M can contain no lattice points. Therefore, under the action of A^{-1}, every lattice point generates a sequence that asymptotically approaches L. This provides a simple way to generate rational approximations to $\sqrt{2}$. Begin with a lattice point $\begin{bmatrix} a \\ b \end{bmatrix}$ and repeatedly apply A^{-1}. Since the resulting sequence of points $\begin{bmatrix} a_n \\ b_n \end{bmatrix}$ approaches L, the ratios a_n/b_n converge to $\sqrt{2}$. For example, starting with $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ we generate the sequence

$$\begin{matrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 3 \\ 2 \end{bmatrix} & \begin{bmatrix} 7 \\ 5 \end{bmatrix} & \begin{bmatrix} 17 \\ 12 \end{bmatrix} \\ \begin{bmatrix} 41 \\ 29 \end{bmatrix} & \begin{bmatrix} 99 \\ 70 \end{bmatrix} & \begin{bmatrix} 239 \\ 169 \end{bmatrix} & \begin{bmatrix} 577 \\ 408 \end{bmatrix} & \begin{bmatrix} 1393 \\ 985 \end{bmatrix} \\ \begin{bmatrix} 3363 \\ 2378 \end{bmatrix} & \begin{bmatrix} 8119 \\ 5741 \end{bmatrix} & \begin{bmatrix} 19601 \\ 13860 \end{bmatrix} & \begin{bmatrix} 47321 \\ 33461 \end{bmatrix} & \begin{bmatrix} 114243 \\ 80782 \end{bmatrix} \end{matrix}$$

The last pair shown approximates $\sqrt{2}$ as $114243/80782$. Squaring the numerator and denominator we find that $114243^2 = 13051463049 = 13051463048 + 1 = 2 \cdot 80782^2 + 1$ so the ratio is indeed very close to $\sqrt{2}$. This same sequence of rational approximations was presented in [**10**, pp. 39–41], derived by an approach closely related to ours, but without using matrices. The sequence also appears in [**8**]. There, a quite different (and very interesting) scheme is used to find rational approximations to $\sqrt{2}$.

Generalizations

The foregoing matrix approach can be generalized in several ways. First we will consider square roots of integers other than 2. Then we will look at the more general case of rational roots of polynomials with coefficients that are either integers, or Gaussian integers. Finally, we generalize from roots (which correspond to linear factors) to the more general question of factorization, as described by Gauss's Lemma.

To begin, let us see how the preceding dynamical discussion of the irrationality of $\sqrt{2}$ generalizes to $\sqrt{n}$. In place of A take the matrix $\begin{bmatrix} -k & n \\ 1 & -k \end{bmatrix}$, and everything works as before. One eigenvalue is $\sqrt{n}-k$ and the corresponding line L of eigenvectors is spanned by $\begin{bmatrix} \sqrt{n} \\ 1 \end{bmatrix}$. The other eigenvalue is $-(k+\sqrt{n})$ with the corresponding line M spanned by $\begin{bmatrix} \sqrt{n} \\ -1 \end{bmatrix}$. In order to obtain the same dynamic behavior as before, we require the first eigenvalue to have magnitude less than 1. We can achieve this by taking k to be the greatest integer in $\sqrt{n}$. In the special case that n is a perfect square, this results in an eigenvalue of 0. Then there *are* lattice points on the line L, but they are all mapped by A to 0 in a single jump. In any other case, that is, if n is not a perfect square, we see that there are no lattice points on L, and deduce that $\sqrt{n}$ is irrational, as before. We have therefore shown that an integer is either a perfect square or has an irrational square root.

One way to view the choice of k in the preceding is as follows: we have a matrix with an eigenvalue that may be larger than 1. By subtracting an integer multiple of the identity matrix, we can translate the eigenvalues to obtain a positive eigenvalue less than 1. This idea leads to a proof of the well known, more general result that a monic polynomial with integer coefficients has real roots that are either integers or irrational. Before proving this result, we need two lemmas. The first allows us to treat a general polynomial in the context of matrix algebra, while the second assures us the equivalent of lattice points as eigenvectors.

LEMMA 1. *Every monic polynomial with integer coefficients is the characteristic polynomial of an integer matrix.*

Proof: The proof is constructive. If the polynomial is $p(x)=x^n+c_{n-1}x^{n-1}+\cdots+c_0$ then it is the characteristic polynomial of the so-called *companion matrix* (see **[5]**, for instance).

$$C=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -c_0 & -c_1 & -c_2 & \cdots & -c_{n-1} \end{bmatrix}$$

It is easy to verify that this matrix has the desired characteristic polynomial by expanding the determinant of $(C-\lambda I)$ in the first column and using induction. Additional insight comes from observing that if α is a root of p, then $[1 \ \alpha \ \alpha^2 \ \cdots \ \alpha^{n-1}]^T$ is an eigenvector of C with eigenvalue α. This fact is easily verified by a direct calculation.

LEMMA 2. *Let A be an integer matrix with rational eigenvalue λ. Then there exists an integer eigenvector u.*

Proof: The matrix $A - \lambda I$ has determinant 0. Therefore, over the field of rationals, it has a nontrivial null space. A nonzero vector in that null space has rational entries, and so a suitable integer multiple will have integer entries. The result is an integer eigenvector u for A and λ.

We state the generalization of the argument concerning $\sqrt{n}$ as follows:

THEOREM 1. *A real eigenvalue of an integer matrix is either an integer or irrational.*

Proof: Proceed by contradiction. Let λ be a rational eigenvalue of the integer matrix A, and assume that λ is not an integer. Without loss, we may assume that $0 < \lambda < 1$, for if not, simply replace A with $A - \lfloor \lambda \rfloor I$. ($\lfloor \cdot \rfloor$ is the greatest-integer function.) The second lemma shows that there is an integer eigenvector u corresponding to λ. If we apply A to u repeatedly, we generate an infinite sequence of distinct integer vectors that converges to 0. This is clearly impossible. Therefore, every rational eigenvalue of A must actually be an integer.

Complex Roots Combined with the first lemma, Theorem 1 shows that for a monic polynomial with integer coefficients any real roots are either integer or irrational. What about the complex roots? To simplify the discussion of the complex case, it will help to use the notation $\mathbb{Z}$ for the integers, and $\mathbb{Z}[i]$ for the Gaussian integers, that is, complex numbers with real and imaginary parts in $\mathbb{Z}$. Similarly, we will denote by $\mathbb{Q}$ the rational numbers, and, by $\mathbb{Q}[i]$, the complex numbers with rational real and imaginary parts. Now let us return to the question of complex roots. If a monic polynomial with integer coefficients has a root in $\mathbb{Q}[i]$, must that root actually lie in $\mathbb{Z}[i]$? The answer is yes, and the argument is essentially the same as what has gone before. Instead of matrices with entries in $\mathbb{Z}$, we consider matrices whose entries are in $\mathbb{Z}[i]$. It is easy to modify the lemmas given earlier to apply to this new situation. First show that every monic polynomial with coefficients in $\mathbb{Z}[i]$ is the characteristic polynomial of a matrix with entries in $\mathbb{Z}[i]$. Then show that when such a matrix has an eigenvalue in $\mathbb{Q}[i]$, it has an eigenvector with entries in $\mathbb{Z}[i]$, as well. Finally, prove that for a Gaussian integer matrix, an eigenvalue in $\mathbb{Q}[i]$ must actually be in $\mathbb{Z}[i]$. As before, it may be assumed without loss of generality that the eigenvalue has magnitude less than 1, this time translating by the nearest Gaussian integer, if necessary. The argument concludes just as before.

Gauss's Lemma All the foregoing results about roots of polynomials can evidently be formulated in terms of linear factors, instead. Thus, if a monic polynomial with integer coefficients has a linear monic factor with a rational constant term, it is actually an integer constant term. This is a special case of a more general result known as Gauss's Lemma: If $f(x)$ is a monic polynomial with integer coefficients which factors as $g(x)h(x)$, where g and h are monic polynomials with rational coefficients, then in fact g and h have integer coefficients. The proof that is usually given for this result makes use of unique factorization. Here, using matrix methods, we can give an alternate proof that does not explicitly depend on unique factorization.

The proof is formulated in terms of *algebraic integers*: complex roots of monic polynomials with integer coefficients. Our preceding results say that an algebraic integer in $\mathbb{Q}$ must be in $\mathbb{Z}$, and an algebraic integer in $\mathbb{Q}[i]$ must be in $\mathbb{Z}[i]$. The first of these results can be applied to prove Gauss's lemma, once we show that the algebraic integers are closed under addition and multiplication. The idea will be to show that the coefficients of factors g and h are algebraic integers since they are combinations of the roots. That will make the coefficients rational algebraic integers, and hence integers.

In addition to its role in the earlier lemmas and results, matrix algebra also provides a convenient means to establish that the algebraic integers are closed under addition and multiplication. It is clear from Lemma 1 that algebraic integers can be characterized as eigenvalues of matrices with integer entries. To deal with sums and products of these eigenvalues, a useful matrix operation is the *tensor product*, also called the *Kronecker product*. Given two matrices A and B, the Kronecker product $A \otimes B$ is defined as follows: Replace each entry a_{ij} of A with an entire block of entries, given by the product $a_{ij}B$. The resulting matrix is $A \otimes B$. There is a nice discussion of Kronecker products in **[6]**. Here, we require only one identity: $(A \otimes B)(C \otimes D) = AC \otimes BD$, which is valid as long as the products AC and BD exist. The proof is a straightforward exercise. With the identity we can prove the following lemma.

LEMMA 3. *If λ and μ are algebraic integers, then so are $\lambda\mu$ and $\lambda + \mu$.*

Proof: Suppose that λ and μ are algebraic integers. Then there are integer matrices A and B, and integer vectors v and w, such that $Av = \lambda v$ and $Bw = \mu w$. Therefore $(A \otimes B)(v \otimes w) = (Av) \otimes (Bw) = \lambda\mu(v \otimes w)$. This shows that $\lambda\mu$ is an eigenvalue of the integer matrix $A \otimes B$, and hence, is an algebraic integer. In a similar way, it is easy to show that $\lambda + \mu$ is an eigenvalue of the integer matrix $A \otimes I + I \otimes B$. Therefore $\lambda + \mu$ is an algebraic integer.

Gauss's lemma is now easily proved.

THEOREM 2. *Let f be a monic polynomial with integer coefficients, and suppose $f = gh$ where g and h are monic polynomials with rational coefficients. Then the coefficients of g and h are actually integers.*

Proof: The roots of f, and hence those of g and h, are algebraic integers. The coefficients of g and h are elementary symmetric functions of the roots, and so can be constructed from the roots using addition and multiplication. This shows that the coefficients of g and h are algebraic integers. But they were assumed to be rational. Thus, they are in fact integers, as asserted.

Integrally Closed Domains We conclude with one further generalization, and a question. The foregoing material can be understood in the context of integral domains and fields of quotients (see, e.g., **[7]**). In our earliest results, the coefficients of the polynomials were integers, and we showed rational roots had to be integers as well. Observe that the rationals are the field of quotients for the integers. This same relationship extends to the results on Gaussian integers. The quotient field for $\mathbb{Z}[i]$ is $\mathbb{Q}[i]$. Our earlier result states that for a monic polynomial over $\mathbb{Z}[i]$, any root in the quotient field of the Gaussian integers must itself be a Gaussian integer.

In both cases, polynomials are considered over an integral domain, and the field of quotients contains no roots other than those that were already present in the integral domain. Proceeding with this more general setting, consider an integral domain D within its field of quotients F. Define $\lambda \in F$ to be *integral over D* if it is a root of a monic polynomial with coefficients in D, and observe that each element of D is integral over D. If these are the only elements integral over D, then D is said to be *integrally closed*. That is, an integral domain D is integrally closed if it contains all the elements of the field of quotients which are integral over D. The earlier results showed that the integers and the Gaussian integers are both integrally closed.

Now the question arises: what is the most general setting for the matrix results presented earlier? Lemmas 1 and 2 still hold if we replace the integers by an integral domain D and the rationals by D's field of quotients. The proofs of Theorem 1 and its extension to the complex case are not so easy to generalize, for they depend on analytic properties that are peculiar to the integers and the Gaussian integers. To illustrate the difficulties, we consider two examples. Each is a quadratic extension of the integers, that is, a domain of the form $\mathbb{Z}[\sqrt{k}\,] = \{n + m\sqrt{k}\,|\,n, m \in \mathbb{Z}\}$ where k is a square-free integer. The field of quotients is $\mathbb{Q}[\sqrt{k}\,]$, defined analogously. It is known that $\mathbb{Z}[\sqrt{k}\,]$ is integrally closed just when $k \not\equiv 1 \pmod 4$. (See, e.g., [7].)

For the first example, $k = -5$, and the domain $\mathbb{Z}[i\sqrt{5}\,]$ is integrally closed. That means that $\lambda \in \mathbb{Q}[i\sqrt{5}\,]$ is a root of a monic polynomial over $\mathbb{Z}[i\sqrt{5}\,]$ only if it is *in* $\mathbb{Z}[i\sqrt{5}\,]$. To demonstrate this, it is tempting to mimic the proof of Theorem 1. Things go awry right at the start, where we want to assume that $|\lambda| < 1$. In the original argument, this step was justified by the observation that λ was at most one unit away from an integer. Unfortunately, that is not true for $\mathbb{Z}[i\sqrt{5}\,]$. Picture the elements as a lattice in the complex plane. The lattice points are separated by one unit horizontally, but by $\sqrt{5}$ units vertically. That means they are too far apart. In particular, if $\lambda = .5 + .5i\sqrt{5}$, the nearest elements of the integral domain are more than one unit away. This foils our desire to find a matrix with entries in $\mathbb{Z}[i\sqrt{5}\,]$ and with an eigenvalue of magnitude less than unity in the quotient field. The argument breaks down because we are unable to produce an appropriate matrix to act as a contraction.

The second example considers $k = 5$, and the result cited earlier says that $\mathbb{Z}[\sqrt{5}\,]$ is not integrally closed. This is easy to see directly: the polynomial $t^2 - t - 1$ has coefficients in $\mathbb{Z}[\sqrt{5}\,]$, and roots $(1 \pm \sqrt{5}\,)/2$ in $\mathbb{Q}[\sqrt{5}\,]$ but not in $\mathbb{Z}[\sqrt{5}\,]$. What happens if we try to follow the proof of Theorem 1 for this example? Observe that all the action takes place on the real line, so the elements of $\mathbb{Q}[\sqrt{5}\,]$ are all within one unit of an integer, and hence, within one unit of an element of the domain $\mathbb{Z}[\sqrt{5}\,]$. As in Theorem 1, we can construct a matrix with an eigenvalue of magnitude less than 1, and which acts as a contraction on the corresponding eigenspace. In particular, a point of that eigenspace with all entries from $\mathbb{Z}[\sqrt{5}\,]$ must generate a sequence of such points converging to the origin. However, for the current example, that presents no contradiction. The elements $\mathbb{Z}[\sqrt{5}\,]$ are not discretely spaced on the real line, and in particular, have 0 for a limit point. So for this example, the entire proof of Theorem 1 remains valid, but failing to result in a contradiction, offers no assurance that $\mathbb{Z}[\sqrt{5}\,]$ is integrally closed.

As these two examples highlight, Theorem 1 and its extension to the Gaussian integers depend on a coincidence of special properties. In addition to the underlying structure exposed in Lemmas 1 and 2, we require a metric on the quotient field satisfying two conditions: (1) the elements of the integral domain cannot get arbitrarily close to 0 (nor hence to any other domain element); and (2) the elements of the domain must get within one unit of every element of the field. In other words, the proof demands that the integral domain elements are neither too close together nor too far apart. This combination of properties does occur for $\mathbb{Z}$ and $\mathbb{Z}[i]$. We don't know if there are any other domains for which the same argument can be made to work, and so we leave it as an open question: Other than $\mathbb{Z}$ and $\mathbb{Z}[i]$, are there integral domains for which the field of quotients satisfies the two conditions above? Clearly, any such domain will have to be integrally closed. That observation prompts another question: Given an integral domain D, under what conditions is there a metric on the field of quotients satisfying the two conditions above?

Conclusion

This paper has considered several aspects of irrationality. Starting with the earliest history, we reviewed the formulation of irrationality in the context of incommensurable segments in geometry. A geometric argument based on infinite descent was reformulated in the now familiar setting of dynamical systems, using matrix algebra for the descent mechanism. In that context, we saw natural extensions from the ring of integers to other structures of modern algebra. In the initial situation, we considered monic polynomials with integer coefficients, and saw that irrational numbers emerge as roots lying outside of $\mathbb{Z}$. The more general setting concerns the monic polynomials over an integral domain D and the nature of roots that are outside of D. The cited result in this area, namely that $\mathbb{Z}[\sqrt{k}\,]$ is integrally closed for square-free k so long as k is not congruent to 1 mod 4, suggests an algebraic subtlety that is absent from the simple dynamic arguments of Theorem 1. Perhaps it should not surprise us that these arguments proved ineffective for $\mathbb{Z}[\sqrt{5}\,]$ and $\mathbb{Z}[i\sqrt{5}\,]$. It remains to be seen whether the dynamic approach can be successfully applied in the more general setting.

REFERENCES

1. Asger Aaboe, *Episodes from the Early History of Mathematics*, New Mathematical Library, Mathematical Association of America, Washington, DC, 1964.
2. Carl Boyer, *A History of Mathematics*, 2nd edition, revised by Uta C. Merzbach, Wiley, New York, NY, 1991.
3. David M. Burton, *The History of Mathematics*, Allyn and Bacon, Newton, MA, 1985, pp. 122–123.
4. James R. Choike, The pentagram and the discovery of an irrational number, *The College Mathematics Journal* 11 (1980), 312–316.
5. I. N. Herstein, *Topics in Algebra*, Ginn and Company, Waltham, MA, 1964.
6. Roger A. Horn and Charles R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.
7. Martin Isaacs, *Algebra, a Graduate Course*, Brooks/Cole, Pacific Grove, CA, 1994, Chapter 28.
8. Thomas Liggett and Peter Peterson, The law of large numbers and $\sqrt{2}$, *The American Mathematical Monthly* 102 (1995), 31–35.
9. Kurt von Fritz, The discovery of incommensurability by Hippasus of Metapontum, *Annals of Math.* 46 (1945), 242–264.
10. Robert Young, *Excursions in Calculus*, Dolciani Mathematical Expositions, Mathematical Association of America, Washington DC, 1992.