Solving Equations, An Elegant Legacy

Jerry L. Kazdan[1]
University of Pennsylvania

Ex irrationalibus oriuntur quantitates impossibiles seu imaginariae,
quarum mira est natura, et tamen non contemnenda utilitas.
[From the irrationals are born the impossible or imaginary quantities
whose nature is very strange but whose usefulness cannot be denied.]
Gottfried Wilhelm Leibniz (1646–1716)

Education is not the filling of a pail, but the lighting of a fire.
William Butler Yeats (1865–1939)

Solving equations. The problems, techniques, and viewpoints are our legacy. One theme throughout this lecture is that classical and modern mathematics are tightly intertwined, that contemporary mathematics contributes real insight and techniques to understand traditional problems.

CONTENTS

In the long second section I discuss some procedures that help to solve equations. I found that the section on symmetry required an extensive discussion because it is treated so inadequately as a fundamental thread throughout mathematics courses. The third section gives three different techniques to prove that equations have solutions. They are typical of those used when an explicit formula for a solution cannot be found.

I organized this so that most of the sections are independent; thus you can skip to examples that are more appealing. To make this more self-contained I have occasionally

[1]This is a considerably expanded version of a lecture—intended for undergraduates—that I gave at both the University of Montreal and the University of Pennsylvania. I thank R. Horn and H. Wilf for valuable suggestions. A shorter version appeared in the *American Math. Monthly*, **105**, Jan. 1998, pp. 1–21.

added details that may not be easily accessible. A few of the tools used here are frequently not met until beginning graduate courses. If these tools are unfamiliar, their appearance here may serve as motivation to learn them.

One ingredient in solving equations that I have not emphasized adequately is the basic role of inequalities. They are lurking here and there: the Euclidean algorithm and the application of the Brouwer fixed point theorem, to name two less obvious instances. It is a shock the first time one sees a proof that $A = B$ not by algebraic manipulation but instead by proving the inequality $|A - B| < \varepsilon$ for any $\varepsilon > 0$. To give inequalities their due would have changed the character of this.

1 Introduction

The simplest equations are of the form

$$2x + 7 = 4.$$

Although the coefficients are positive integers, one is forced to enlarge the type of possible solution to include not only rational numbers, but also negative numbers. It took centuries for negative numbers to be accepted. Through the Middle Ages they frequently were called *false* numbers.

The next sort of equation one meets is perhaps

$$x^2 = 2.$$

Again to solve this one must enlarge the type of possible solution to include the *irrational* number $\sqrt{2}$. The word *irrational* itself reveals people's emotional attitudes. Another word used for numbers such as $\sqrt{2}$ is *surd*, which is related to the word "absurd."

The equation

$$x^2 + 1 = 0$$

again forces one to introduce new types of numbers, the *imaginary* numbers. The quotation from Leibniz at the beginning of this article conveys the views of his era.

These complex numbers were adequate to solve all quadratic equations

$$ax^2 + bx + c = 0. \tag{1}$$

From the explicit formula $(-b \pm \sqrt{b^2 - 4ac})/2a$ for the solutions x_1 and x_2, one observes that

$$x_1 + x_2 = -\frac{b}{a} \quad \text{and} \quad x_1 x_2 = \frac{c}{a}. \tag{2}$$

For further progress it was essential that one also could obtain these formulas *without* using the explicit formula for the solution. One merely expands

$$0 = (x - x_1)(x - x_2) = x^2 - (x_1 + x_2)x + x_1 x_2 \tag{3}$$

and compares the coefficients with those of (1). This was an early significant instance where one found properties of the solutions of an equation without first requiring a formula for the solution.

After using complex numbers to solve quadratic equations, it was, however, surprising that complex numbers were also adequate to find a formula to solve the general cubic polynomial equation $p(x) := ax^3 + bx^2 + cx + d = 0$. One does not need to enlarge further beyond the complex numbers. Without using the formula for the roots it is obvious how to obtain the analog of (2); if the roots are x_1, x_2, x_3, then expanding $p(x) = a(x - x_1)(x - x_2)(x - x_3)$ we get for instance

$$x_1 + x_2 + x_3 = -\frac{b}{a} \tag{4}$$

An immediate consequence is that if the coefficients in the polynomial are rational and if two of the roots are rational, then so is the third root.

Eventually, an explicit formula for the solutions of a quartic equation was also found. Here too, complex numbers were adequate to find all solutions. In the seventeenth century there was probably uncertainty if $\sqrt{i}$ was a complex number. That one could write $\sqrt{i} = \pm(1+i)/\sqrt{2}$ would have surprised many—including Leibniz.

Solving the general quintic polynomial was a challenge. If the coefficients of

$$p(x) := x^5 + bx^4 + cx^3 + dx^2 + ex + f. \tag{5}$$

are real, obviously for all large positive x we have $p(x) > 0$, while for all large negative x we have $p(x) < 0$. Thus if you graph the polynomial $y = p(x)$, it is geometrically evident that it crosses the x-axis at least once and hence there is at least one real root x_1 of $p(x) = 0$. The polynomial $q(x) := p(x)/(x - x_1)$ is then a quartic polynomial for whose four roots there are formulas. Thus it was known that every quintic polynomial has five (some possibly repeated or complex) roots. It was upsetting when Abel [1802–29] showed that despite knowing these five roots *exist*, there cannot be a general *formula* for them that involves only the usual arithmetic operations along with taking roots. Formulas similar to (4) were essential in Abel's reasoning.

Mathematicians found themselves in the fascinating dilemma of having proved that these roots exist but also having proved that there can never be an algebraic formula for them. The general existence proof is what we now call the *Fundamental Theorem of Algebra*, while understanding the obstructions to finding formulas for the roots is *Galois* [1811–1832] *theory*. Both were vital pillars in the future development of mathematics. As a twist of fate, except for their fundamental historic role, the formulas for the solutions of the cubic and quartic have become museum pieces, rarely used because they are so complicated.

The proof that the quintic (5) always has at least one real root was one of the first "pure" existence proofs. Although this proof was regarded as obvious, in the nineteenth century mathematicians became more concerned because this proof presumes that the real number line has no "holes." What would happen if there were a hole in the number line exactly where the root should have been? How can one precisely define this "no holes" property?

After considerable effort, mathematicians learned how to make precise what they meant when they said that the number line has no "holes." Ever since, the resulting concept, *completeness,* has been a basic ingredient in mathematics. One reason that it is so important to consider the class of all Lebesgue [1875–1941] integrable functions is that by including them the function spaces L^p are complete.

By allowing polynomials to have complex roots, one can prove that a polynomial of degree n has exactly n roots—if one counts multiple roots appropriately. The number of real roots is considerably more complicated and depends on the coefficients of the polynomial (Sturm's theorem [Wf]). This is why when one studies the roots of simultaneous polynomial equations, which is the focus of algebraic geometry, one usually uses a field, such as the complex numbers, where polynomials of degree n have exactly n roots. Not much is known about polynomials if one works only with the real numbers.

2 Steps Toward Solving Equations

In solving equations, the most primitive question is to decide if there are any solutions at all. From our understanding of the special case of polynomial equations, we have learned to separate this from the important problem of explicitly finding solutions. Moreover, in the many cases where we know there is a solution but there is no "formula", you need qualitative properties of the solution.

2.1 **What does "solution" mean?**

It may be necessary to broaden what an acceptable solution is, much as for polynomials we usually allow complex solutions, perhaps in projective space. You may solve a diophantine equation mod p for all primes p. For partial differential equations one accepts solutions in various function spaces, including distribution and Sobolev spaces of functions. Finding the appropriate notion of "solution" may be a key step.

2.2 **Find a formula for a solution**.

Usually there is no formula of any sort. Even when is there is one, it may involve a reduction to another problem, say finding the roots of a polynomial or evaluating an integral, which you accept as a solution. But this acceptance depends on the personal background of the consumer. In earlier centuries difficulties were faced if the "solution" of a problem involved numbers like $\sqrt{7}$ or π, or, worse yet, complex numbers. Similarly, many people have difficulty accepting a power or Fourier [1768–1830] series as the solution of any equation. For them infinite series are problems, not answers. From the power series for $\sin x$, the 2π periodicity is far from obvious; that information is more accessible from other approaches. Eventually, one learns that even an infinite series solution may encode useful information, although it takes experience before one learns to find them useful.

A numerical solution may be valuable in some circumstances, yet in others it may be a jumble of numbers that you need to decipher to learn anything useful. Hamming's assertion: "The purpose of computing is insight, not numbers," applies to most scientific computations.

There are elementary problems where there is no formula for the solution, but there is an algorithm for finding a solution. Even in such cases occasionally you may prefer a non-constructive proof that a solution exists.

An example is solving $ax \equiv b \pmod{m}$, where a and m are relatively prime. Since the solution is $x \equiv a^{-1}b \pmod{m}$, we need to find $a^{-1} \pmod{m}$. One traditional approach is to observe that the numbers $a, 2a, \ldots, (m-1)a$ are all distinct $\pmod{m}$ so one of them must be $1 \pmod{m}$. This proof that a^{-1} exists gives no hint of how to find it except by trial and error. This is a non-constructive existence proof for the solution of $ax \equiv 1 \pmod{m}$. One constructive proof considers the equivalent problem of solving $ax - my = 1$ for integers x, y. The Euclidean algorithm solves this explicitly (see [Da, Section I.8]). Since at the k^{th} step in this algorithm the absolute value of the remainder can be chosen to be at most half the value of the previous remainder, this new remainder is at most $a/2^k$ so you need at most $\log a/\log 2$ steps (this is one of the few places that we consider the important issue of the efficiency of an algorithm).

An alternative approach to find a^{-1} is to use the Fermat [1601–65]–Euler [1707–83] identity $a^{\varphi(m)} \equiv 1 \pmod{m}$, where the Euler function $\varphi(m)$ is the number of integers k, with $1 \le k \le m-1$ that are relatively prime to m (if $m = p$ is a prime number then $\varphi(p) = p-1$). Thus $a^{-1} \equiv a^{\varphi(m)-1} \pmod{m}$. Note, however, that computing $a^{\varphi(m)-1} \pmod{m}$ requires as much calculation as exhaustively testing $a, 2a, \ldots, (m-1)a$; the method using the Euclidean algorithm is faster.

Polynomial interpolation supplies an example where a variety of approaches are available to solve some equations, each approach with its own illumination. Here we seek a polynomial $p(x) := a_0 + a_1 x + \cdots + a_k x^k$ of degree k with the property that its graph $y = p(x)$ passes through $k+1$ specified points $(x_1, y_1), \ldots, (x_{k+1}, y_{k+1})$, where the x_j's are distinct. Thus we can view the problem as solving the $k+1$ linear equations $p(x_j) = y_j$, $j = 1, \ldots, k+1$ for the $k+1$ coefficients $a_0, \ldots, a_k$.

Method 1 (Lagrange [1736–1813]). Lagrange introduced a clever basis for the space $\mathcal{P}_k$ of polynomial of degree at most k. It is $e_j(x) = \Pi_{i \ne j}\big[(x - x_i)/(x_j - x_i)\big]$, $j = 1, \ldots, k+1$. Thus $e_j(x_i) = \delta_{ij}$, the Kronecker delta (this was perhaps the first instance in mathematics of a "dual basis"). Then the explicit—unique—solution to the interpolation problem is simply

$$p(x) = y_1 e_1(x) + y_2 e_2(x) + \cdots + y_{k+1} e_{k+1}(x). \tag{6}$$

Method 2 (Newton [1642–1727]). Seek $p(x)$ in the special form

$$p(x) = A_0 + A_1(x - x_1) + A_2(x - x_1)(x - x_2) + \ \cdots \ + A_k(x - x_1)\cdots(x - x_k).$$

Setting $x = x_1$ we find $y_1 = p(x_1) = A_0$. Then set $x = x_2$ to find A_1, and so on, finding succeeding coefficients $A_2, \ldots, A_k$ recursively. Here we use Newton's basis 1, $(x - x_1)$, $(x - x_1)(x - x_2)$, $\ldots$, $(x - x_1)(x - x_2)\cdots(x - x_k)$ for $\mathcal{P}_k$.

Method 3. Define the linear map $L : \mathcal{P}_k \to \mathbb{R}^{k+1}$ by the rule

$$L : p \mapsto (p(x_1), p(x_2), \ldots, p(x_{k+1})).$$

The interpolation problem is to find the inverse map of L. Observe that if $Lp = 0$, then $p(x) \in \mathcal{P}_k$ vanishes at the $k+1$ distinct points $x_1, \ldots, x_{k+1}$, therefore p must be the zero polynomial. Thus, the kernel of L is zero. Since both $\mathcal{P}_k$ and $\mathbb{R}^{k+1}$ have the same dimension, $k+1$, then by basic linear algebra the map L is invertible. This proves that the interpolation problem has a unique solution—but yields no formulas or special procedures.

Comparing these, Method 1 yields some basic information quickly, but is not easy to use to compute $p(x)$ at other points (too many multiplications). The formula for Method 2 is computationally much easier to use to evaluate $p(x)$. It has two additional virtues. (i) If the polynomial $p(x)$ is an approximation to some other function, $f(x)$, then you can use this method to find an estimate for the error $|f(x) - p(x)|$. This error estimate is similar to that found for Taylor series (see any book on numerical analysis, my favorite is [D-B, p. 100]). (ii) If you add another interpolation point x_{k+2}, then the formulas for the coefficients A_j already computed do not change. Finally, Method 3 shows quickly that the problem has a unique solution. See our discussion of harmonic polynomials in Section 2.5 for a less obvious application of Method 3.

We'll give a brief application of interpolation to numerical integration. Say you want to evaluate $J := \int_a^b f(x)\,dx$. You specify $k+1$ distinct points $a \le x_1 < \ldots < x_{k+1} \le b$ and seek a formula

$$\int_a^b f(x)\,dx \approx B_1 f(x_1) + B_2 f(x_2) + \cdots + B_{k+1} f(x_{k+1}). \tag{7}$$

Can one find coefficients B_j so this formula is exact whenever f happens to be a polynomial of degree at most k? Yes, and the B_j's are unique. A naive approach is to let f be x^j, $j = 0,..,k$ in (7). This gives $k+1$ linear equations for the $k+1$ unknowns $B_1,\ldots,B_{k+1}$. But it is simpler to use the Lagrange formula (6) to find the polynomial interpolating f at the chosen points: $f(x) \approx p(x) = \sum_{j=1}^{k+1} f(x_j)e_j(x)$. Then

$$\int_a^b f(x)\,dx \approx \int_a^b p(x)\,dx = \sum_{j=1}^{k+1} f(x_j) \int_a^b e_j(x)\,dx.$$

Thus $B_j = \int_a^b e_j(x)\,dx$. The trapezoidal rule is the special case of the two points $x_1 = a$, $x_2 = b$, while Simpson's rule is for the three points, $x_1 = a$, $x_2 = (a+b)/2$, $x_3 = b$.

Finally we might also want to pick the points x_j themselves so the formula (7) is exact for all polynomials of even higher degree. Since the right side of (7) now involves $2k+2$ parameters (the B_j's and x_j's), we suspect by choosing them adroitly we should be able to have (7) be exact for all polynomials of degree $2k+1$ (its dimension is $2k+2$). The equations (7) are linear in the B_j's but nonlinear in the x_j's. Gauss studied this. He found that the x_j's should be chosen as the zeroes of the polynomial φ_{k+1} of degree $k+1$ from the family of orthogonal polynomials associated with the standard inner product $\langle u, v\rangle = \int_a^b uv\,dx$ (for $a=-1$, $b=1$ these are the Legendre polynomials). We prove Gauss' result.

The proof is short and clever. Let Q be any polynomial of degree at most $2k+1$. By division write $Q = q\varphi_{k+1} + r$, where q and r are polynomials of degree at most k. The orthogonality property gives $\langle q, \varphi_{k+1}\rangle = 0$; combined with our choice of the B_i's to make the formula exact for all polynomials of degree k we find

$$\int_a^b Q(x)\,dx = \int_a^b r(x)\,dx = \sum_{j=1}^{k+1} B_j r(x_j). \tag{8}$$

But since the x_j's are zeroes of φ_{k+1} then

$$\sum_{j=1}^{k+1} B_j Q(x_j) = \sum_{j=1}^{k+1} B_j[\varphi_{k+1}(x_j) + r(x_j)] = \sum_{j=1}^{k+1} B_j r(x_j). \tag{9}$$

Comparing (8) and (9) we see that the integration procedure (7) is exact for Q.

2.3 **Find an equivalent problem that is simpler.**

a) *Change of variable*

Making a *change of variable* is perhaps the most familiar technique to simplify a problem. A small example of this is the cubic polynomial $p(x) = ax^3 + bx^2 + cx + d$. View the coefficients as those in a Taylor series. Since the second derivative is zero at the point where $6ax + 2b = 0$, the change of variables $z = 6ax + 2b$ (or just the translation $z = x + b/3a$) yields a simpler polynomial $q(z) = \alpha z^3 + \gamma z + \delta$ without a quadratic term. If the coefficients of the original equation were rational, then so are those of the new equation and the rational roots of the new equation correspond to those of the original equation. This is a generalization of the procedure of "completing the square." Similarly, by a translation one can eliminate the coefficient a_{n-1} in $p(x) = x^n + a_{n-1}x^{n-1} +$ lower order terms

We can use this to show that *every double root of a cubic polynomial with rational coefficients is rational*. Using our change of variable, it is enough to show this for $q(z) = \alpha z^3 + \gamma z + \delta$. Thus, we must show that if $q(r) = 0$ and $q'(r) = 0$, then r is rational. But $0 = q'(r) = 3\alpha r^2 + \gamma$ implies that $\alpha r^3 = -(\gamma/3)r$. Thus $0 = q(r) = -(\gamma/3)r + \gamma r + \delta$, that is, $r = -3\delta/2\gamma$, which is rational. From (4), since $x_1 = x_2 = r$, the third root of q (and hence of p) is also rational.

For cubic polynomials with rational coefficients and having a double root r (necessarily rational, from the above) you can now find all *rational points* (x, y) (that is, both x and y are rational) on the "elliptic curve" $y^2 = p(x)$. They are the points where straight lines through $(r, 0)$ and having rational slope intersect the curve. This is now an easy exercise. A related exercise is to show that the rational points on the circle $x^2 + y^2 = 1$ are where the straight lines through $(1, 0)$ with rational slope intersect the circle. One consequence is a formula for all the "Pythagorean triples": the integers a, b, c with $a^2 + b^2 = c^2$.

Another instance of finding an equivalent problem that is simpler is the change of variable (that is, a change of basis) in a matrix equation to diagonalize the matrix (if possible). We can use the same idea for a system of differential equations

$$Lu := u' + Au = f, \tag{10}$$

where $u(t)$ and $f(t)$ are vectors and $A(t)$ is a square matrix. We seek a change of variables $u = Sv$ where $S(t)$ is an invertible matrix, to transform this to a simpler equation. In some applications this is called a *gauge transformation.* To find a useful S we compute

$$f = Lu = u' + Au = (Sv)' + A(Sv) = Sv' + (S' + AS)v. \tag{11}$$

The right side of this is simplest if S is a solution of the matrix equation

$$LS = S' + AS = 0, \qquad \text{say with} \qquad S(0) = I; \tag{12}$$

we use $S(0) = I$ to insure that S is invertible. Then solving (11) is just integrating $v' = g$ where $g = S^{-1}f$.

With this choice of S and writing $D := d/dt$ it is instructive to rewrite (11) as $f = Lu = SDv = SDS^{-1}u$. In particular, $L = SDS^{-1}$. One sees that *every* linear ordinary differential operator is "conjugate" or "gauge equivalent" to D. We thus come to the possibly surprising conclusion that *any* first order linear differential operator L is equivalent to the simple operator D; this makes studying linear ordinary differential operators far easier than partial differential operators. We also have formally $L^{-1} = SD^{-1}S^{-1}$. Since D^{-1} is integration (and adding a constant of integration), an immediate consequence is that the general solution of the inhomogeneous equation $Lu = f$ is

$$u(t) = L^{-1}f = S(t)C + S(t)\int_0^t S^{-1}(\tau)f(\tau)\,d\tau, \tag{13}$$

where $C = u(0)$. The matrix S defined by (12) is the usual *fundamental matrix solution* one meets for ordinary differential equations. Unfortunately it is presented frequently as a trick to solve the inhomogeneous equation rather than as a straightforward approach to reduce the study of L to the simpler differential operator D. It is sometimes useful to introduce *Green's function* (G. Green [1793–1841]) $G(t,\tau) := S(t)S^{-1}(\tau)$ and rewrite (13) as

$$u(t) = u(0) + \int_0^t G(t,\tau)f(\tau)\,d\tau. \tag{14}$$

We then think of the integral operator with kernel $G(t,\tau)$ as L^{-1}. This integral can be interpreted physically and gives another (equivalent) approach to solving (10).

Usually S cannot be found explicitly. However in special cases such as a single equation or a 2×2 system with constant coefficients, you can carry out the computations and obtain the classical formulas quickly. For instance, for a single equation, we find that $S(t) = e^{-\int A(t)\,dt}$. Then we recognize (13) as the standard formula. Since one can write a second order linear ODE as a first order system of this form, we have also covered that case.

What we call a "change of variable" is part of a fundamental procedure known to everyone, yet often seems exotic when it arises in a mathematical setting. As an illustration, say you have a problem $\mathcal{P}$ that is stated in another language, perhaps Latin. To solve it, first translate (T) it into your language, solve the translated version $\mathcal{Q}$, and then translate it back (T^{-1}). Symbolically, reading from *right* to *left*,

$$\mathcal{P} = T^{-1}\mathcal{Q}T$$

$$\begin{tikzcd} \textit{old} \arrow[r, "\mathcal{P}", "\text{(original version)}"'] \arrow[d, "T"'] & \textit{old} \\ \textit{new} \arrow[r, "\mathcal{Q}", "\text{(new version)}"'] & \textit{new} \arrow[u, "T^{-1}"'] \end{tikzcd}$$

Figure 1

(see Figure 1). The goal is to choose the new setting and T so the new problem $\mathcal{Q}$ is easier than $\mathcal{P}$. Diagonalizing a matrix and using a Laplace [1749–1827] transform are two familiar mathematical examples. The same idea—but with a different twist—is also useful in discussing symmetry (Section 2.7). There we will see that finding a T so that the new version is the *same* as the old, $\mathcal{P} = T^{-1}\mathcal{P}T$, is how one identifies symmetries of the problem $\mathcal{P}$. As a silly linguistic illustration, one observes the phrase "madam I'm Adam" reads the same backwards. Here T is the operation of reading backward.

b) *Variational problem*

The *calculus of variations* offers a radical way to reformulate some problems. First an example in $\mathbb{R}^n$. If A is a self-adjoint matrix, then solving $Ax = b$ is equivalent to finding a critical point of the scalar-valued function

$$J(x) = \tfrac{1}{2}\langle Ax, x\rangle - \langle x, b\rangle,$$

where we use the standard inner product in $\mathbb{R}^n$. To see this, if x is a critical point of $J(x)$ and if we set $\varphi(\varepsilon) = J(x + \varepsilon v)$, where v is a vector and ε a scalar, then by the definition of critical point, $\varphi'(0) = 0$. But by a computation $\varphi'(0) = \langle Ax - b, v\rangle$. Thus $\langle Ax - b, v\rangle = 0$ for *any* vector v. Since v is arbitrary this implies that $Ax - b = 0$, as we asserted.

A related problem is to find the lowest eigenvalue of a self-adjoint matrix A. It is an simple exercise to show this is the minimum value of the function $J(x) = \langle x, Ax\rangle / \|x\|^2$ as x ranges over all vectors $x \neq 0$ (or, equivalently, minimize $\langle x, Ax\rangle$ on the unit sphere $\|x\| = 1$). A vector x giving this minimum value is a corresponding eigenvector. This approach to the eigenvalues of a self-adjoint matrix is used widely in computations, as well as giving one way to prove that one can always diagonalize a self-adjoint matrix. Since it does not use the fundamental theorem of algebra, this approach is applicable to some problems in infinite dimensional spaces—such as Sturm ([1803–55]–Liouville ([1809–82]) theory and the spectrum of the Laplacian (see just below).

The identical approach works for more complicated problems. Say we want to solve the *wave equation*, $u_{tt} = u_{xx} + u_{yy}$ to find the position $u(x,y,t)$ of a vibrating membrane Ω; thus (x,y) are in a region $\Omega \in \mathbb{R}^2$ and time t is a real number. We claim that u being a solution of the wave equation is equivalent to u being a critical point of the functional

$$J(u) = \tfrac{1}{2}\iint_{\Omega}\int_{\alpha}^{\beta}(u_t^2 - u_x^2 - u_y^2)\,dx\,dy\,dt. \tag{15}$$

One verifies this formally just as in the previous example by considering $\varphi(\varepsilon) = J(u + \varepsilon v)$, where v is any smooth function with compact support in $\Omega \times (\alpha, \beta)$. Again, by definition of critical point this means $\varphi'(0) = 0$ for any of our functions v. By differentiating under the integral

$$\varphi'(0) = \left.\frac{dJ(u + \varepsilon v)}{d\varepsilon}\right|_{\varepsilon=0} = \iint_{\Omega}\int_{\alpha}^{\beta}(u_t v_t - u_x v_x - u_y v_y)\,dx\,dy\,dt.$$

To simplify this, we integrate by parts (the divergence theorem), taking the derivative off the v terms and placing them on the u terms. There are no boundary terms because we assumed v had compact support. The previous equation reads

$$\varphi'(0) = -\iint_{\Omega}\int_{\alpha}^{\beta}[u_{tt} - u_{xx} - u_{yy}]\,v\,dx\,dy\,dt.$$

From this it is clear that the solutions of the wave equation are critical points of φ. For this converse, it is helpful to introduce the inner product $\langle f, g\rangle = \iiint fg\,dx\,dy\,dt$. Then since $\varphi'(0) = 0$,, the last formula asserts that the expression in brackets $[\cdots]$ is orthogonal to all these functions v. Since smooth functions v with compact support are dense, the expression in brackets must be zero. That is, u must be a solution of the wave equation. It is customary to refer to the wave equation as the *Euler-Lagrange equation* for the functional $J(u) = \int F(x,t,u,u_t,u_x,u_y)\,dx\,dy\,dt$ whose integrand $F(x,t,u,u_t,u_x,u_y) := \frac{1}{2}(u_t^2 - u_x^2 - u_y^2)$ is called the *Lagrangian*.

Closely related to the linear algebra case, the lowest eigenvalue of the Laplacian, $-\Delta u = \lambda u$ (note the "$-$" sign) for functions with zero boundary values on a region Ω is found by minimizing the Rayleigh [1842–1919] quotient $J(v) = \int_{\Omega}|\nabla v|^2\,dx / \int_{\Omega} v^2\,dx$ among all functions v that are zero on the boundary of Ω. This is useful both for theoretical and practical applications.

One virtue of introducing a variational problem is that some properties may be more accessible. We see instances of this below, where we'll use invariance of the variational problem under the translation $t \mapsto t+\varepsilon$ to deduce conservation of energy for the wave equation (Section 2.7d), and in a situation where the existence of a solution to the original problem is more accessible from the variational approach (Section 3.2). Two standard references to the calculus of variations are [G-F] (a basic text) and [G-H] (a fresh, more thorough, approach). The book [H-T] is a nice introduction for the general reader.

2.4 **Duality: Find a related problem that is useful.**

To me, duality is the most vague and mysterious item in this lecture. My impression is that duality appeared first in projective geometry where one interchanges the roles of points and lines (this evolved slowly from Apollonius' [*c.* 262–190 B.C.] use of "pole" and "polar" through 1850). Lagrange introduced the adjoint of a differential operator in the eighteenth century (this is the essence of Lagrange's identity for linear second order ordinary differential operators) while the adjoint of a matrix seems to have been used significantly only in the nineteenth century. Green's second identity (1828) asserts that the Laplacian is formally self-adjoint. Lagrangian and Hamiltonian mechanics are dual objects: Lagrangian living on the tangent bundle, Hamiltonian on the cotangent bundle. There are dual problems in the calculus of variations—including linear programming. Cohomology is the dual of homology. Duality is even a standard device in rhetoric: "Do unto others as you would want others do unto you", and J.F. Kennedy's "... ask not what your country can do for you, ask what you can do for your country". I do not know how to make the concept of duality precise enough to fit all known mathematical instances and ease introduction of new dual objects.

In Section 2.6 below we give a more subtle uses of duality in linear algebra and differential equations. As preparation, and for its own interest, here we follow Lagrange and define the formal adjoint L^* of a linear differential operator L. Use the inner product for real-valued functions: $\langle \varphi, \psi \rangle = \int \varphi\psi\, dx$. Then L^* is defined by the usual rule

$$\langle u, L^* v\rangle = \langle Lu, v\rangle$$

for all smooth functions u and v that are zero outside a compact set; we choose functions that are zero outside a compact set to avoid having boundary terms when we integrate by parts. We use the word "formal" since the strict adjoint requires a (complete) Hilbert [1862–1943] space and the consideration of boundary conditions.

If $L := d/dt$, then an integration by parts reveals that

$$\langle Lu, v\rangle = \int u' v\, dt = -\int uv'\, dt = \langle u, L^* v\rangle.$$

Thus, the formal adjoint of $L := d/dt$ is $L^* = -d/dt$. Similarly, if $A(t)$ is a matrix and $u(t)$ is a vector, then the formal adjoint of $Lu := u' + A(t)u$ is $L^* v = -v' + A^*(t)v$. Two integrations by parts show that the formal adjoint of the second order system $Mu := u'' + A(t)u$ is $M^* v = v'' + A^*(t)v$. In particular, if A is symmetric then M is formally self-adjoint, a fact that is basic in quantum mechanics, where, with a complex inner product, the self-adjoint operator id/dt appears in the Schrödinger equation.

One application of the adjoint is that if u is a solution of the homogeneous system $Lu := u' + A(t)u = 0$ and v is a solution of the adjoint system, $L^* v = -v' + A^*(t)v = 0$, then their pointwise inner product $v \cdot u$ is a constant. Indeed,

$$\frac{d}{dt}(v \cdot u) = v' \cdot u + v \cdot u' = A^* v \cdot u - v \cdot Au = 0. \tag{16}$$

Observing that $v \cdot u$ is the matrix product $v^* u$, a similar computation shows that if $S(t)$ and $T(t)$ are (not necessarily square) matrix solutions of $LS = 0$ and $L^* T = 0$, respectively, then

$$T^*(t)S(t) = \text{constant}. \tag{17}$$

In particular, if A, S and T are square matrices with $S(0) = T(0) = I$ (as in (12), S and T are then *fundamental matrix solutions*), we have

$$T^*(t)S(t) = I \qquad \text{that is,} \qquad T(t) = S^{-1*}(t). \tag{18}$$

If this formula (17) appears boring, the disguise is perfect. It is a wide-sweeping generalization both of $e^t e^{-t} = 1$, which is the special case of $Lu := u' + u$, so $L^* v = -v' + v$, as well as $\cos^2 t + \sin^2 t = 1$. In a physical context it may express some conservation law.

To prove $\cos^2 t + \sin^2 t = 1$, consider the second order system $Mw := w'' + Cw = 0$, where $C(t)$ is an $n \times n$ matrix, with corresponding adjoint system $M^* z = z'' + C^* z$. Let $\varphi(t)$ and $\psi(t)$ be (vector or matrix) solutions of $M\varphi = 0$ and $M^* \psi = 0$, respectively. We assert that

$$\psi^{*\prime}(t)\varphi(t) - \psi^*(t)\varphi'(t) = \text{constant}. \tag{19}$$

This reduces to $\cos^2 t + \sin^2 t = 1$ in the special case where C is the 1×1 identity matrix, $\varphi(t) = \cos t$, and $\psi(t) = \sin t$. The identity (19) is a routine consequence of the basic identity (17) and requires no additional insight to discover; merely rewrite $w'' + C(t)w = 0$ as a first order system by the usual procedure of letting $u_1 := w$ and $u_2 := w'$. Then $u := \binom{\varphi}{\varphi'}$ satisfies the first order system $Lu := u' + Au$, where A is the $2n \times 2n$ block matrix $A := \begin{pmatrix} 0 & -I \\ C & 0 \end{pmatrix}$. Similarly $v := \binom{\psi'}{-\psi}$ is a solution of the adjoint equation $L^* v = -v' + A^* v = 0$. The result (19) now follows from the identity (17) with $S = u$ and $T = v$.

For equations of the form $Lu := Pu' + Au$, where P may be singular at a boundary point of the interval under discussion (this arises in Sturm-Liouville theory), it is useful to observe that probably one should *not* multiply by P^{-1} to reduce to the earlier case. Instead directly use $L^* v = -(P^* v)' + A^* v$ and generalize the identity (17) to $T^*(t)P(t)S(t) = \text{const.}$ Similarly, for $Mw := (Pw')' + Cw = 0$ identity (19) becomes $\psi^{*\prime} P\varphi - \psi^* P\varphi' = \text{const.}$

A consequence of (18) and (14) is that if $G(t,\tau) = S(t)S^{-1}(\tau)$ is Green's function for $Lu := u' + Au$, then Green's function for L^* is $G^*(\tau, t)$, a fact that has the useful physical interpretation that for the adjoint one interchanges the roles of the observation time t and event time τ (to see this clearly for a scalar equation let $f(\tau)$ be the Dirac delta function at, say, $\tau = \tau_0$ in (14)).

2.5 **Understand the family of all solutions**.

How many solutions are there? Is uniqueness desirable? If so, what conditions would insure uniqueness of the solution? If you slightly modify some parameters in the problem, do the solutions change only slightly? This continuous dependence on parameters is basic in real-life problems where the data are known only approximately. It is also important for problems solved using a computer that introduces both round-off errors (computers use only a finite number of decimal places) and truncation errors (computers approximate limiting processes such as integration by finite discrete operations).

For instance, by Rouche's theorem in complex analysis the roots of a polynomial $p(z)$ depend continuously on the coefficients, that is, if p has k roots in the small disk $|z-c|<\rho$ and if we perturb the coefficients of p slightly, then this perturbed polynomial also has exactly k roots in this disk. A corollary is that the eigenvalues of a matrix depend continuously on the elements of the matrix. The simple example $x^2=\pm\varepsilon$ shows that these assertions may be *false* if one considers only *real* roots.

This example $x^2=\varepsilon$ for ε near zero also shows that, even allowing complex roots, the solution may *not* be a differentiable function of the parameter. By contrast, we will use the implicit function theorem to show easily that simple roots do depend smoothly on parameters. Here is the proof. Say we have a polynomial $p(x,c)$ depending smoothly on a parameter c and at $c=c_0$ we have a root x_0, so $p(x_0,c_0)=0$. Since x_0 is a simple root, $\partial p(x,c_0)/\partial x|_{x=x_0}\neq 0$. Thus, by the implicit function theorem, for all c near c_0, there is a unique solution $x=x(c)$ near x_0 of $p(x,c)=0$. This solution depends smoothly on c. This proof was quite general; it does not require p to be a polynomial. A consequence is that simple eigenvalues of a matrix depend smoothly on the elements of the matrix.

The study of what happens when one cannot apply the implicit function theorem is carried out in *bifurcation theory* and in the study of *singularities of maps*. We now know these are the same subjects, although they arose from different origins with different viewpoints. See [Arn], [C-H], [H-K], and [G-S]. The key new phenomenon is that several solutions can branch—or solutions can disappear—as occurs for the real solutions of $x^2=\varepsilon$; for $\varepsilon<0$ there are no real solutions while for $\varepsilon>0$ there are two solutions. An early notable appearance of bifurcation theory was Euler's classical study of the buckling of a slender column under compression (see the elementary discussion in the undergraduate text [Wi, pp. 167-169]).

In practical problems, one may need to delve more deeply into the dependence of a problem on parameters. Wilkinson (see Forsythe's beautifully illuminating article [F] and subsequent book [F-M-M]) illustrated this with the polynomial

$$p(x)=(x-1)(x-2)\cdots(x-19)(x-20)=x^{20}-210x^{19}+\cdots .$$

Let $p(x,\varepsilon)$ be the polynomial obtained by replacing only the term $-210x^{19}$ by $-(210+\varepsilon)x^{19}$, where $\varepsilon=2^{-23}$. Since $2^7<210<2^8$, this means we are changing this one coefficient in the 30^{th} significant base 2 digit. A smaller perturbation of this sort might even occur because of roundoff error in a computer, because computers keep only a finite number of decimal places. Since for $\varepsilon=0$ the roots of $p(x,0)$ are well-separated, then for ε near zero they depend smoothly on ε. By a careful calculation, one finds that some of the roots have moved substantially. For instance the complex numbers $16.73073\pm 2.81262i$ are now roots. Should we be surprised the roots have moved this much? No. For if we differentiate $p(x,\varepsilon)=0$ with respect to ε we obtain

$$\frac{\partial p(x,\varepsilon)}{\partial x}\frac{\partial x}{\partial \varepsilon}+\frac{\partial p(x,\varepsilon)}{\partial \varepsilon}=0,$$

so

$$\frac{\partial x}{\partial \varepsilon} = -\frac{\partial p/\partial \varepsilon}{\partial p/\partial x} = \frac{x^{19}}{\sum\limits_{1\le j\le 20} \prod\limits_{\substack{1\le k\le 20\\ k\ne j}} (x-k)}.$$

Evaluating this at $x = j$ for $j = 1, \ldots, 20$ we find the sensitivity of the j^{th} root:

$$\left.\frac{\partial x}{\partial \varepsilon}\right|_{x=j} = \frac{j^{19}}{\prod\limits_{\substack{1\le k\le 20\\ k\ne j}} (j-k)}.$$

For instance, at the root $x = 16$ one computes that $\left.\frac{\partial x}{\partial \varepsilon}\right|_{x=16} = 2.4 \times 10^9$. Not small at all.

One can explicitly find the family of all solutions for only the simplest problems, yet these frequently serve as guides for more general cases. To the astute—at least with hindsight—the polynomial equation $z^n = 1$ and differential equation $u'' = f(x)$ give significant hints of how more complicated cases behave. Even without explicit formulas, you can sometimes obtain information on the set of all possible solutions. The following example illustrates this; it also is an instructive indication of the power of Method 3 in Section 2.2.

Consider the linear space $\mathbb{P}_\ell$ of polynomials of degree at most ℓ in the n variables $x_1, \ldots, x_n$ and let P_ℓ be the sub-space of polynomials homogeneous of degree ℓ. The standard Laplacian on $\mathbb{R}^n$ is $\Delta u := u_{x_1x_1} + u_{x_2x_2} + \cdots u_{x_nx_n}$. A function $u(x)$ is called *harmonic* if $\Delta u = 0$. We wish to compute the dimension of the subspace H_ℓ of P_ℓ consisting of homogeneous harmonic polynomials. If $n = 2$, and $\ell \ge 1$ the dimension is 2, since for $\ell \ge 1$ one basis for the space of harmonic polynomials of degree exactly ℓ is the real and imaginary parts of the analytic function $(x+iy)^\ell$.

For the general case, observe that $\Delta : P_{\ell+2} \to P_\ell$ and define the linear map $L : \mathbb{P}_\ell \to \mathbb{P}_\ell$ by the formula

$$Lp(x) := \Delta\big[(|x|^2 - 1)p(x)\big], \tag{20}$$

where $|x|$ is the euclidean norm. Now $Lp = 0$ means the polynomial $u(x) := (|x|^2 - 1)p(x) \in \mathbb{P}_{\ell+2}$ is harmonic. But clearly $u(x) = 0$ on the sphere $|x| = 1$, so $u \equiv 0$.[2] Thus $\ker L = 0$ so L is invertible. In particular, given a homogeneous $q \in P_\ell$ there is a $p \in \mathbb{P}_\ell$ with $\Delta\big[(|x|^2 - 1)p(x)\big] = q$. Let $v \in P_\ell$ denote the homogeneous part of p that has highest degree ℓ. Since Δ reduces the degree by two, we deduce that in fact $\Delta(|x|^2 v) = q$. Therefore this map $v \mapsto q$ from $P_\ell \to P_\ell$ is onto and hence an isomorphism.[3] Here are two consequences.

1) Since the map $\Delta : P_\ell \to P_{\ell-2}$ is onto, again by linear algebra, we can compute the dimension of the space of homogeneous harmonic polynomials:

$$\dim H_\ell = \dim P_\ell - \dim P_{\ell-2} = \binom{n+\ell-1}{\ell} - \binom{n+\ell-3}{\ell-2} = \frac{(n+2\ell-2)(n+\ell-3)!}{\ell!(n-2)!}.$$

[2]To prove $u \equiv 0$, one can use the divergence theorem to see that $\int_{|x|<1} |\nabla u|^2\, dx = -\int_{|x|<1} u\Delta u\, dx = 0$, so $\nabla u = 0$. Thus $u \equiv \text{const.} \equiv 0$. Another approach uses the maximum principle for harmonic functions.

[3]One can also give a purely algebraic proof that if $p \in P_\ell$ satisfies $\Delta(|x|^2 p) = 0$, then $p \equiv 0$ —hence the map $M : P_\ell \mapsto P_\ell$ defined by $Mp := \Delta(|x|^2 p)$ is an isomorphism of P_ℓ.

For instance if $n=3$ then $\dim H_\ell = 2\ell+1$.
2) Any homogeneous polynomial $q \in P_\ell$ can be written (uniquely) in the form $q = h + |x|^2 v$, where $h \in H_\ell$ and $v \in P_{\ell-2}$. To prove this, first compute Δq and then use the above to find a unique $v \in P_{\ell-2}$ so that $\Delta(|x|^2 v) = \Delta q \in P_{\ell-2}$. The function $h := q - |x|^2 v$ is clearly harmonic. Applying this again to v and so on recursively we conclude that $q = h_\ell + |x|^2 h_{\ell-2} + |x|^4 h_{\ell-4} + \cdots$, where $h_j \in H_j$. This yields the direct sum decomposition $P_\ell = H_\ell \oplus |x|^2 H_{\ell-2} \oplus \cdots$. Since both the Laplacian and the operation of multiplying by $|x|^2$ commute with rotations (see the discussion in Section 2.7a below), the summands in this decomposition are $SO(n)$-invariant, a fact that is useful in discussing spherical harmonics and the symmetry group $SO(n)$.

The idea behind the definition of L in (20) was that to solve $\Delta u = q \in \mathbb{P}_\ell$, we seek u in the special form $u = (|x|^2-1)p(x)$ to obtain a new problem, $Lp = q$, whose solution is unique. Frequently it is easier to solve a problem if you restrict the form of the solution to obtain uniqueness.

Homogeneous harmonic polynomials arise since, when restricted to the unit sphere you can show that they are exactly the eigenfunctions of the Laplacian on the sphere; the dimensions of the eigenspaces are the numbers just computed. As above, when $n=3$ this number is $2\ell+1$. Atoms are roughly spherically symmetric and this number arises as the maximum number of electrons in an atomic subshell. There are $2\ell+1$ electrons with spin $\pm\frac{1}{2}$, so $2(2\ell+1)$ in all. Thus the subshells contain at most $2, 6, 10, 14, \ldots$ electrons.

In high school we solve polynomial equations in one variable and systems of linear equations. These are the first steps in understanding the solutions of a system of k polynomial equations

$$f_1(z) = 0,\ f_2(z) = 0, \ldots,\ f_k(z) = 0$$

in n unknowns $z = (z_1, \ldots, z_n)$. From experience we know that it is simplest to allow complex numbers as solutions. If there are more equations than unknowns ($k > n$), then usually there will be no solutions, that is, no common zeroes [CHALLENGE: restate this precisely and then prove it, say for any smooth functions f_j], while if there are more unknowns than equations there are usually infinitely many solutions. If there are the same number of equations as unknowns, then usually there are only finitely many solutions. While plausible, this is not obvious (it is false for non-polynomials as $\sin x = 0$, which has infinitely many solutions—hardly a surprise if one views its Taylor series as a polynomial of infinite degree).

Bezout [1730–83] made this precise for two polynomial equations in two variables:

$$f(x,y) = 0 \qquad g(x,y) = 0. \tag{21}$$

If f has degree k and g degree ℓ, he proved that there are exactly $k\ell$ solutions, possibly complex, unless f and g have a common (non-constant) polynomial factor (see [Wa], [Ful]).

As usual, it is enlightening to introduce geometric language and think of the two equations (21) as defining two curves C_1 and C_2. The common solutions of (21) are the points where the curves intersect.

We examine (21) in the special case where $f(x,y) = a_1x^3 + a_2x^2y + \cdots + a_{10}$ and g are both cubic polynomials. Say they intersect at the nine points $p_1 = (x_1, y_1), \ldots, p_9 =$

(x_9, y_9). So far this is quite general. But now assume that six of these, $p_1, \ldots, p_6$, happen to lie on a conic Γ, so they are also roots of the quadratic polynomial $q(x,y) = 0$ that defines Γ. By Bezout, we know that C_1 and Γ intersect in six points, so it is quite special if C_2 and Γ intersect in the *same* six points. For simplicity also assume that the conic Γ is irreducible, that is, it is not the product of two non-constant polynomials of lower degree (this is the case if Γ is the product of two linear polynomials and hence is just two straight lines). We claim that the remaining three points p_7, p_8, p_9 lie on a straight line.

Here is an algebraic proof. For any linear combination $h(x,y) := \alpha f(x,y) + \beta g(x,y)$, notice that the cubic curve C defined by $h = 0$ automatically contains the points where C_1 and C_2 intersect. Pick another point v on the conic Γ and choose α and β so that v is also a zero of h. Then the cubic curve C also intersects the conic Γ at the *seven* points $v, p_1, \ldots, p_6$. But by Bezout's theorem C and Γ have $3 \cdot 2 = 6$ points of intersection unless h and q have a common factor. Thus there must be a common factor. Because q is irreducible, the factor must be q itself, so $h(x,y) = q(x,y)r(x,y)$ where, by matching degrees, $r(x,y)$ is a linear polynomial. Thus p_7, p_8, p_9, which are zeroes of $h = 0$ but not $g = 0$, are roots of the linear polynomial $r = 0$ and thus lie on a straight line.

We can reinterpret this to obtain a classical theorem of Pascal [1623–1662]. Connect any six points $p_1, \ldots, p_6$ on a conic to obtain a "hexagon", probably with self-intersections. Some terminology for hexagons: a pair of sides separated by two sides is called *opposite* (as p_1p_2 and p_4p_5) while the points of intersection of opposite sides are called *diagonal points*. Thus a hexagon has three diagonal points (circled in Figure 2). Pascal's theorem asserts that these three points always lie on a straight line.

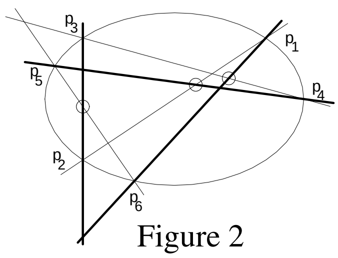

Figure 2

To prove it, take the alternate edges of the hexagon, $\overline{p_1p_2}$, $\overline{p_3p_4}$, $\overline{p_5p_6}$, and $\overline{p_2p_3}$, $\overline{p_4p_5}$, $\overline{p_6p_1}$, to obtain two triangles whose sides contain these edges. To each triangle we associate a cubic polynomial by taking the product of the three linear polynomials determined by the edges of the triangle. Note that here a triangle is the union of the three entire lines, not just the segments joining vertices. Then the points $p_1, \ldots, p_6$ plus the three diagonal points are the nine points of intersection of these triangles. Now apply the preceding algebraic result. To include the possibility that some pairs of opposite sides might be parallel—so the corresponding points of intersection are at infinity—it is better if one works in the projective plane.

The algebraic reasoning generalizes immediately: *Let $f(x,y) = 0$, $g(x,y) = 0$ be polynomials of degree n that intersect at n^2 points. If kn of these points lie on an irreducible curve defined by a polynomial of degree k, then the remaining $n(n-k)$ points lie on a curve defined by a polynomial of degree $n-k$.* This generalization illustrates the power of the algebraic approach, despite the loss of the special beauty of a purely synthetic geometric proof.

2.6 If a solution does not always exist, find the obstructions.

If an equation does not have a solution, it is important to understand the reason. If you are trying to fit a straight line $p = at + b$ to the k data points $(t_1, p_1), \ldots, (t_k, p_k)$ that were found experimentally, then it is unlikely there will be a choice of the coefficients a and b that fits the data exactly. In this situation one seeks an "optimal" approximate solution. A typical approach to solving $F(x) = y$ approximately is to find a solution x_0 that minimizes the error: $E(x) = \|F(x) - y\|$. A non-trivial human decision is choosing a norm (or some other metric) for measuring the error. One often uses a norm arising from an inner product; the procedure is then called the *Method of Least Squares*. Here is a brief (but complete) outline.

First observe that if a linear space V has an inner product (written as $\langle x, y\rangle$), and $S \subset V$ is a subspace, then the *orthogonal projection* y_S of y into S has the property that $y - y_S$ is perpendicular to S. *The projection* y_S *is the point in* S *closest to* y since for any $w \in S$ we have $y - w = (y - y_S) + (y_S - w)$. Because $(y - y_S) \perp (y_S - w) \in S$, by Pythagoras

$$\|y - w\|^2 = \|y - y_S\|^2 + \|y_S - w\|^2 \geq \|y - y_S\|^2.$$

For least squares to minimize the error $\|Lx - y\|$ we thus want to pick x so that $y_S := Lx$ is the orthogonal projection of y into $S := \text{image}\,(L)$. Then $y - y_S = y - Lx$ will be perpendicular to $\text{image}\,(L)$ so for every vector z

$$0 = \langle y - Lx, Lz\rangle = \langle L^*(y - Lx), z\rangle.$$

Therefore $L^*(y - Lx) = 0$. We can rewrite this by saying the desired x is a solution of the *normal equation* $L^*Lx = L^*y$. Since L^*L is a square matrix, the normal equations have a (unique) solution if $\ker(L) = 0$. [If you have never done so, a simple but useful exercise is to set-up the normal equations for the above example of fitting a straight line to some data.]

There are situations where other procedures are more appropriate to minimize the error $F(x) - y$. For nonlinear problems not much is known. In linear and nonlinear programming there is related work to find optimal solutions of *inequalities*.

Now, say you want to solve an equation that you believe *should* have an exact solution under suitable conditions. You thus need to determine and understand these conditions.

The simplest case is a system of linear algebraic equations $Ax = y$ (the matrix A is not assumed to be square). A basic—but insufficiently well known—result in linear algebra uses the adjoint equation (duality) and says that for a given y there is at least one solution if and only if y is orthogonal to all the solutions z of the homogeneous adjoint equation, $A^*z = 0$. Here is the short proof.

Say z satisfies $A * z = 0$. If there is a solution of $Ax = y$, then taking the inner product with z we obtain

$$\langle z, y\rangle = \langle z, Ax\rangle = \langle A^*z, x\rangle = 0. \tag{22}$$

Thus z is orthogonal to y. This computation shows that

$$\text{image}\,(A)^\perp = \ker(A^*). \tag{23}$$

The proof of formula (23) is the same in infinite dimensional Hilbert spaces. If V is a linear subspace of a Hilbert space, then $(V^\perp)^\perp = \overline{V} = \text{closure of } V$. Thus, in a Hilbert space, $\overline{\text{image}(A)} = \ker(A^*)^\perp$. In some important cases—including $\mathbb{R}^n$ where it is evident—one can show that $\text{image}\,(A)$ is closed. One then has

$$\text{image}\,(A) = \ker(A^*)^\perp. \tag{24}$$

Frequently this is called the *Fredholm* [1866–1927] *alternative*, since it can be phrased as the following alternative: "Either you can always solve $Ax = y$, or else there are obstructions. These obstructions are precisely that y must be orthogonal to all the solutions of the homogeneous adjoint equation."

As another example, consider solving the differential equation $u'' = f$, where we assume $f(x)$ is periodic, say with period 2π, and we seek a solution $u(x)$ that is also periodic with the same period, so both u and u' are periodic (that u'' will be periodic follows from the differential equation). Thus we are solving the differential equation on the circle, S^1. This is a simple example of an "elliptic differential operator" on a "compact manifold without boundary."

First we solve the equation directly. The general solution of $u'' = f$ is $u(x) = u(0) + u'(0)x + \int_0^x (x-t)f(t)\,dt$. To insure that u is periodic we need $u(2\pi) = u(0)$ and $u'(2\pi) = u'(0)$. The second condition imposes the requirement

$$\int_0^{2\pi} f(x)\,dx = 0, \tag{25}$$

and we use the first condition to solve for $u'(0)$. The upshot is that a solution exists if and only f satisfies (25). This solution is not unique since the constant $u(0)$ can be chosen arbitrarily.

Next we interpret (25) using the Fredholm alternative. Write our equation as $Lu = f$, where $Lu := u''$, so L is formally self-adjoint: $L^*v = v''$. The Fredholm alternative says that to find the image of L, we should first find the periodic solutions of the homogeneous adjoint equation, $z'' = 0$. Although this equation can be solved by a mental computation, we use a different method that generalizes. Multiply the equation $z'' = 0$ by z and integrate by parts to obtain

$$0 = \langle z, Lz\rangle = \int_0^{2\pi} zz''\,dx = -\int_0^{2\pi} |z'|^2\,dx, \tag{26}$$

so $z' = 0$ and z is constant. The Fredholm alternative then states that f is in the image of L precisely when it is orthogonal to the constants:

$$\langle 1, f\rangle = 0.$$

This is just equation (25).

This example may be generalized to solving the Laplace equation on the torus T^n. Here we are given a (smooth) function $f(x_1, x_2, \ldots, x_n)$ that is periodic with period 2π in each variable and seek a periodic solution $u(x_1, x_2, \ldots, x_n)$ of

$$\Delta u = f(x), \tag{27}$$

where $\Delta u := u_{x_1x_1} + u_{x_2x_2} + \cdots + u_{x_nx_n}$. Using Fourier series, it is straightforward to show that there is a solution if and only if the same condition holds,

$$\int_{T^n} f(x)\,dx = 0, \tag{28}$$

where, in this formula we integrate from 0 to 2π in each variable $x_1, x_2, \ldots, x_n$. Just as for $u'' = f$ the Fredholm alternative gives (28) for the Laplace equation (27); to compute $\ker\Delta$ you merely replaces the integration by parts in (26) by the divergence theorem (cf the footnote in Section 2.5).

Almost 100 years ago, Fredholm proved that the Fredholm alternative holds for the Laplace equation. We now know that it holds for many linear "elliptic" partial differential equations with various boundary conditions. The Fredholm alternative is more interesting for these differential operators than in finite dimensional spaces since for them the kernels of L and L^* are finite dimensional (this is elementary for ordinary differential operators, but deeper for elliptic partial differential operators). Thus there are only a finite number of obstructions to solving $Lu = f$, despite the function space being infinite dimensional.

The Hodge [1903–75] theorem for compact manifolds is a straightforward consequence (essentially algebraic) of the Fredholm alternative applied to the Hodge Laplacian on differential forms.

Since it is both instructive and (to my surprise) not readily accessible in the literature, we will show in detail that the Fredholm alternative holds for the second order ordinary differential equation

$$Mu := a(x)u'' + b(x)u' + c(x)u = g(x), \tag{29}$$

where the coefficients and $g(x)$ and their derivatives are smooth functions that are periodic with period 2π. To avoid singularities also assume $a(x) \neq 0$. We seek a smooth solution $u(x)$ that is also periodic with period 2π. In other words, we are solving (29) on the circle $S^1 = \{0 \leq x \leq 2\pi\}$ with the end points $x = 0$ and $x = 2\pi$ being thought of as the same point.

The computation (22) shows that if g is in the image of M, that is, if you can solve (29), then g is orthogonal to the kernel of M^*. The converse is more complicated.

The details are a bit simpler if we assume we have already made the standard reduction to a first order system of the form

$$Lu := u' + A(x)u = f(x), \tag{30}$$

where $A(x)$ is a square matrix and $f(x)$ a vector, with all the elements of A and f being smooth periodic functions (we always assume the period is 2π). We seek a smooth periodic (vector) solution u. Our short proof uses the existence theorem for ordinary differential equations.

One tool we use is the *fundamental matrix solution* $S(x)$ and the resulting formula (13) for the general solution of the inhomogeneous equation (knowing one can find $S(x)$ is the only place we use the existence theorem for ordinary differential equations). The question thus reduces to finding a constant vector $C := u(0)$ so that u is periodic, that is, $u(2\pi) = u(0)$. Using (13) we can write $u(2\pi) = u(0)$ as

$$[I - S(2\pi)]C = S(2\pi)\int_0^{2\pi} S^{-1}(t)f(t)\,dt. \tag{31}$$

From (31) it is clear that in the special case where 1 is *not* an eigenvalue of $S(2\pi)$, that is, if the homogeneous equation has no solutions with period 2π, then we can solve (31) uniquely for C. But 1 might be an eigenvalue of $S(2\pi)$; we must look deeper.

Both to treat the general case and to relate this to the homogeneous adjoint equation $L^*v = -v' + A^*(x)v = 0$, we need the observation (18) that the fundamental matrix solution of the adjoint operator is S^{*-1}. Thus the general solution (not necessarily periodic) of $L^*z = 0$ is $z(t) := S^{*-1}(t)Z$ where Z can be any vector. Consequently $z(t)$, which we have just noted is a solution of the homogeneous adjoint equation, is periodic with period 2π if and only if $S^{-1*}(2\pi)Z = S^{-1*}(0)Z$, that is, if $Z \in \ker[S^{-1*}(2\pi) - I]$.

From here, the reasoning is straightforward. For instance we deduce the Fredholm alternative for periodic solutions of (30) follows. Rewrite (31) as

$$[S^{-1}(2\pi) - I]C = \int_0^{2\pi} S^{-1}(t)f(t)\,dt.$$

Let V be the right hand side of this. By linear algebra, one can solve this algebraic equation for C if and only if V is orthogonal to $\ker[S^{-1}(2\pi) - I]^*$, that is, to all vectors $Z \in \ker[S^{-1*}(2\pi) - I]$. However, V being orthogonal to these vectors Z means

$$0 = Z \cdot V = \int_0^{2\pi} Z \cdot S^{-1}(t)f(t)\,dt = \int_0^{2\pi} z(t) \cdot f(t)\,dt,$$

where $Z \cdot V$ is the usual inner product in $\mathbb{R}^n$ Consequently, (30) has a periodic solution if and only if f is orthogonal in $L_2(S^1)$ to the periodic solutions of the homogeneous adjoint equation. This completes the proof.

Another easy consequence of this approach is that the dimension of the space of 2π periodic solutions of $Lu = 0$ and of $L^*v = 0$ are equal. Indeed, from (31), the dimension of the space of periodic solutions of $Lu = 0$ is $\dim\ker[I - S(2\pi)]$. Similarly, since S^{*-1} is the fundamental matrix for L^*, then the dimension of the space of periodic solutions of $L^*v = 0$ is $\dim\ker[I - S^{*-1}(2\pi)]$. But $I - S^{*-1} = -[(I - S)S^{-1}]^*$. Thus $I - S$ is just $I - S^{*-1}$ multiplied by an invertible matrix and then taking an adjoint so the dimensions of their kernels are equal.

The reader may wish to use these ideas to prove that the Fredholm alternative holds for the boundary value problem $L := u'' + c(x)u = f(x)$ on the interval $0 < x < 1$, with the "Dirichlet" boundary conditions $u(0) = 0$, $u(1) = 0$ for both L and L^*.

All of this has treated linear equations. Understanding obstructions to existence for nonlinear equations is much more complicated, even in Euclidean space for real solutions of a system of polynomial equations.

The next example gives the flavor of the issues for a simple nonlinear differential equation. Recall that the curvature $k(x)$ of a smooth curve $y = y(x)$ is given by

$$k(x) = \frac{y''}{(1 + y'^2)^{3/2}}. \tag{32}$$

The "inverse curvature problem" is, given a smooth function $k(x)$, $0 < x < 1$, to find a smooth curve $y = y(x)$ having this function as its curvature.

A circle of radius R has curvature $1/R$. Thus, if $k(x) \equiv 2$, then a semi-circle of radius $1/2$ solves our problem. However, if $k(x) \equiv 4$, then the circle of radius $1/4$ supplies a solution for only half the desired interval $0 < x < 1$. This leads us to suspect that if there is a solution, then the curvature can't be too large for too much of the interval.

To find an obstruction, note that $y''/(1+y'^2)^{3/2} = (y'/\sqrt{1+y'^2})'$. Thus we integrate both sides of (32)

$$\int_0^x k(t)\,dt = \frac{y'(x)}{\sqrt{1+y'(x)^2}} - \frac{y'(0)}{\sqrt{1+y'(0)^2}}. \tag{33}$$

Let $\gamma = y'(0)/\sqrt{1+y'(0)^2}$, so $|\gamma| \le 1$ and

$$\int_0^x k(t)\,dt \le 1-\gamma \le 2, \qquad 0 \le x \le 1.$$

This inequality embodies our suspicion that "the curvature can't be too large for too much of the interval." For the case of constant curvature $k(x) \equiv c > 0$, for $x = 1$ this condition is $c \le 2$, which is sharp. For non-constant k a necessary and sufficient condition is that there is a constant $\gamma \in [-1,1]$ such that $\left|\int_0^x k(t)\,dt + \gamma\right| < 1$ for all $0 < x < 1$. If we assume the curve is convex, that is, $k(x) > 0$, then we may choose $\gamma = -1$ and find that a necessary and sufficient condition is simply $\int_0^1 k(t)\,dt \le 2$. The necessity is immediate from (33), while the sufficiency follows by solving (33) for $y'(x)$ and integrating. Implicitly we have not permitted vertical tangents ($y'(x) = \pm\infty$) inside the interval but do allow them at the boundary points—as in the case of a semicircle of radius $1/2$.

A standard variant of this problem is to impose *boundary conditions* such as $y(0) = y(1) = 0$. I leave you the pleasure of discovering necessary and sufficient conditions for solving this boundary value problem in the special case of a convex curve. Assuming existence, is the solution of this boundary value problem unique?

Another variant: For a plane curve $(x(s), y(s))$ parameterized by arc length, $0 \le s \le L$, one can compute the curvature $k(s)$. Investigate the inverse problem: given $k(s)$, $0 \le s \le L$, find the curve. What if you require the curve to be a smooth (simple?) closed curve?

The difficulties here are because this problem is *global* for the whole interval $0 < x < 1$. If we are satisfied with a *local* solution, defined only in some neighborhood of $x = 0$ then a solution always exists.

For surfaces $z := u(x,y)$ in $\mathbb{R}^3$ one analogue of this uses the *mean curvature* H:

$$H(x,y) = \nabla\cdot\Big(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) \tag{34}$$

For instance the mean curvature of a sphere of radius R is $2/R$, while the mean curvature of a right circular cylinder of radius R is $1/R$. For a cylindrical surface where $z = u(x)$ does not depend on y, the mean curvature equals the curvature of the curve $z = u(x)$.

The inverse mean curvature problem is, "Given $H(x,y)$ and a connected region $\Omega \in \mathbb{R}^2$, is there a surface $z = u(x,y)$ having mean curvature H for all $(x,y) \in \Omega$?"

As in the previous case, we anticipate that if H is too large in some sense, then the desired surface will not exist over all of Ω. The obstruction is a (possibly surprisingly) straightforward extension of (33). Integrate both sides of (34) over any region $\omega \subseteq \Omega$ with sufficiently smooth boundary $\partial\omega$. Then by the divergence theorem

$$\iint_\omega H(x,y)\,dx\,dy = \int_{\partial\omega} \frac{\nabla u\cdot\nu}{\sqrt{1+|\nabla u|^2}}\,ds,$$

where ds is the element of arc length and ν the unit outer normal vector field. Since $|\nabla u\cdot\nu|/\sqrt{1+|\nabla u|^2}\le 1$ we have the obstruction

$$\left|\iint_\omega H(x,y)\;dx\,dy\right| \le \text{Length}\,(\partial\omega).$$

In particular, if $H(x,y)\ge$ const. $c>0$ and Ω is a disk of radius R, then $c\le 2/R$ (see [K, p. 37] for a bit more).

Our understanding of obstructions to the existence of a solution of most nonlinear partial differential equation is very incomplete; many of the known obstructions use Noether's theorem mentioned in Section 2.7d. The border between existence and non-existence is still largely uncharted territory.

2.7 **Exploit symmetry.**

a) *Simple symmetry*

One familiar example of symmetry in algebra occurs for a polynomial $p(z)=a_nz^n+\cdots+a_0$ with real coefficients. Here the coefficients are invariant under complex conjugation so for any complex number z we have $\overline{p(z)}=\sum\overline{a_kz^k}=\sum a_k\bar{z}^k=p(\bar{z})$. Thus if z is a complex root, then so is $\bar{z}$. Since taking the complex conjugate a second time brings us back to the original root, we don't get even more roots this way (but in the last example in this section, repeatedly using a symmetry will give us infinitely many integer solutions of $x^2-2y^2=1$). The nature of complex conjugation as a symmetry is clearer if one uses different (more cumbersome) notation for the complex conjugation operator, say write $T(z)=\bar{z}$. Thus $T^2=$Identity and $(Tp)(z)=T(p(z))$. For a polynomial with real coefficients $\overline{p(z)}=p(\bar{z})$ means $Tp=pT$, that is, T and p commute; it may be clearer if we write this as $TpT^{-1}=p$, so p is fixed under the automorphism T. Galois' deep contribution to the theory of solving polynomial equations was to show how to exploit related symmetries.

A variant of this reasoning is also useful to solve the equation $F(x)=c$. Assume that F commutes with some map T, so $TF=FT$, and that c is invariant under T: $T(c)=c$. If x_0 is a solution of $F(x)=c$, then x_0 is not necessarily invariant, but by the above reasoning $T(x_0)$ is also a solution. If you also know that the solution of $F(x)=c$ is *unique*, then $T(x_0)=x_0$, that is, this solution x_0 is invariant under T. Here are three similar instances.

i). Let f be a homeomorphism of the sphere $S^2\in\mathbb{R}^3$, and let $\varphi:(x,y,z)\mapsto(x,y,-z)$ be a reflection across the equator. Assume that $c\in S^2$ is fixed by φ, that is, $\varphi(c)=c$ so c is on the equator $z=0$, and assume that f and φ commute, $f\circ\varphi=\varphi\circ f$. If $f(p_0)=c$, then $p_0=(x_0,y_0,z_0)$ is also invariant under φ and hence p_0 is also on the equator. Thus f maps the equator onto itself.

ii). The second example is the solution $u(x,t)$ of the wave equation $u_{xx}-u_{tt}=0$ on the interval $-1\le x\le 1$ with the boundary conditions $u(-1,t)=u(1,t)=0$. If the initial position $u(x,0)$ and the initial velocity $u_t(x,0)$ are both even functions, that is, invariant under the map $T : x \mapsto -x$, then so is the solution $u(x,t)$. This follows as soon as you know the uniqueness of the solution of the wave equation with given initial conditions. [4]

Using the linearity of this problem, even if we did not have uniqueness we could still have obtained an invariant solution by letting $\varphi(x,t)$ be any solution; since $T^2 = I$, then the average $u := \frac{1}{2}(\varphi + T\varphi)$ is an invariant solution. One generalizes the construction of u in similar situations by the important procedure of averaging over the group of symmetries. One application in electrostatics is the *method of images*.

iii). A Markov chain example. In an experiment you are placed in a five room "house" (see Figure 3). Every hour the doors are opened and you must move from your current room to one of the adjacent rooms. Assuming the rooms are all equally attractive, what percentage of the time will you spend in each room? (The extent to which the experimental percentage differs from this measures the desirability of each room).

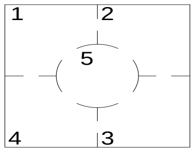

Figure 3

To solve this problem one introduces the 5×5 *transition matrix* $M = (m_{ij})$ of this *Markov* [1856–1922] *chain*: if you are currently in room j, then m_{ij} is the probability you will next be in room i (CAUTION: some mathematicians interchange the roles of i and j). For this, we number the rooms, say clockwise beginning in the upper left corner with p_5 referring to the center room. Then, for instance, $m_{12} = m_{32} = m_{52} = \frac{1}{3}$ since if you are in room 2, it is equally likely that you will next be in rooms 1, 3, or 5, but you won't be in rooms 2 or 4. Proceeding similarly we obtain

$$M = \begin{pmatrix} 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{4} \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}$$

The elements of M are non-negative and the sum of every column is 1: no matter where you are now, at the next step you will certainly be in one of the rooms.

It is useful to introduce column *probability vectors* $P = (p_1, \ldots, p_5)$ with the property that p_j gives the probability of being in the j^{th} room at a given time. Then $0 \le p_j \le 1$ and $\sum p_j = 1$. If P_{now} describes the probabilities of your current location, then $P_{\text{next}} = MP_{\text{now}}$, gives the probabilities of your location at the next time interval. Thus, if one begins in Room 1, then $P_0 = (1,0,0,0,0)$, and after the first hour $P_1 = (0, \frac{1}{3}, 0, \frac{1}{3}, \frac{1}{3}) = MP_0$. In the same way, at the end of the second hour $P_2 := MP_1 = M^2 P_0$, and $P_k := MP_{k-1} = M^k P_0$.

For a matrix M arising in a Markov process (non-negative elements and the sum of each column is one), if λ is any eigenvalue of M^* (and hence M), then $|\lambda| \le 1$. To see this, let $v := (v_1, \ldots, v_n)$ be a corresponding eigenvector, $M^* v = \lambda v$, with largest component v_k, that is, $|v_i| \le |v_k|$. Then $|(\lambda - m_{kk})v_k| = |\sum_{i \ne k} m_{ik} v_i| \le (\sum_{i \ne k} m_{ik})|v_k|$. Since $\sum_i m_{ik} = 1$ then $|\lambda - m_{kk}| \le 1 - m_{kk}$. Consequently $|\lambda| \le |\lambda - m_{kk}| + m_{kk} \le 1$ (this reasoning is a special case of Gershgorin's theorem).

---

[4] Proof of Uniqueness. Say u and v are both solutions with the same initial position and velocity, then $w := u - v$ is also a solution with $w(x,0) = w_t(x,0) = 0$. Apply conservation of energy (45) to $w(x,t)$. Since $E(0) = 0$, then $E(t) \equiv 0$ for all t. Hence $w(x,t) \equiv$ const.. Since $w(x,0) = 0$, then $w(x,t) \equiv 0$ so $u(x,t) \equiv v(x,t)$.

Moreover, if we assume all the elements of M are positive, then equality $|\lambda| = 1$ occurs only if $\lambda = 1$ and $v_1 = v_2 = \ldots = v_n$. Thus $|\lambda| < 1$ except for the one dimensional eigenspace corresponding to $\lambda = 1$.

In seeking the long-term probabilities, we are asking if the probability vectors $P_k = M^k P_0$, $k = 1, 2, \ldots$ converge to some "equilibrium" vector P *independent* of the initial probability vector P_0. If so, then in particular $P = \lim M^{k+1} P_0 = \lim M M^k P_0 = MP$, that is, $P = MP$ so P is an eigenvector of M with eigenvalue 1. Moreover, choosing P_0 to be *any* standard basis vector e_j and since the j^{th} column of M^n is $M^n e_j \to P$, it follows that $M^k \to M_\infty$ where all the columns of M_∞ are the *same* eigenvector P. In addition, still assuming convergence to equilibrium, every eigenvector of M with eigenvalue $\lambda = 1$ must be a multiple of P.

Although $\lambda = 1$ is always an eigenvalue of M (since it is an eigenvalue of M^* with eigenvector $(1, \ldots, 1)$), the limit $M^k P_0$ does not always exist. For example, it does not exist for the transition matrix $M = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ for a two room "house." If $M = I$, then the limit of $M^k P_0$ exists but is not independent of P_0. However the limit $M^k P_0$ does exist and is independent of the initial probability vector P_0 if all of the elements of M —or some power of M—are positive. If M is diagonalizable, this follows from the above information on its eigenvalues. For the general case one must work harder. [5] In our case all the elements of M^2 are positive since after two steps there is a positive probability that one will be in each of the rooms. It remains to find this limiting probability distribution P by solving $P = MP$.

Here is where we can use symmetry. Since the four corner rooms are identical, M must commute with the matrices T_{ij} that interchange the probabilities of being in the corner rooms, p_i and p_j for $1 \le i, j \le 4$. Since $M(T_{ij}P) = T_{ij}MP = T_{ij}P$, we see that $T_{ij}P$ is also a probability eigenvector with eigenvalue $\lambda = 1$. Thus, by uniqueness of this probability eigenvector, $T_{ij}P = P$ so "by symmetry" P has the special form $P = (x, x, x, x, y)$ with $1 = \sum p_i = 4x + y$. The system of equations $P = MP$ now involves only two unknowns x, y. Its first equation is $x = \frac{1}{3}x + \frac{1}{3}x + \frac{1}{4}y$, that is $4x = 3y$. Combined with $4x + y = 1$ one finds $x = \frac{3}{16}$, $y = \frac{1}{4}$. Therefore 25% of the time is spent in the center room and 18.75% in each of the corner rooms. Symmetry turned a potentially messy computation into a simple one.

---

[5] The simplest proof I know for the convergence without assuming M is diagonalizable is in [Be, p. 257]. One shows that M^{*k} converges as $k \to \infty$ to a matrix M^*_∞ each of whose rows are the same, so for any given column all the elements are the same. Since the proof does not seem to be widely known, here is a sketch. AVERAGING LEMMA: *If one takes a weighted average $\overline{w} = c_1 w_1 + c_2 w_2 + \cdots + c_n w_n$ of real numbers $w_1, \ldots, w_n$, where $0 < \gamma \le c_j$ and $c_1 + \cdots + c_n = 1$, then the average lies between the max and min of the w_j with the quantitative estimate $\gamma w_{\max} + (1-\gamma) w_{\min} \le \overline{w} \le (1-\gamma) w_{\max} + \gamma w_{\min}$.*

To apply this let $\gamma > 0$ be the smallest element of M. Because the sum of the elements in any row of M^* is 1, if w is any vector then the elements of $z := M^* w$ are various averages of w. Thus the above estimate gives the upper bound for $z_{\max} \le (1-\gamma) w_{\max} + \gamma w_{\min}$ and similarly $\gamma w_{\max} + (1-\gamma) w_{\min} \le z_{\min}$. These imply $z_{\max} - z_{\min} \le (1 - 2\gamma)(w_{\max} - w_{\min})$. Because $0 < 1 - 2\gamma < 1$, iterating this contraction proves that each element of the vector $M^{*k} w$ converges to the *same* number. To get the j^{th} column of M^*_∞ use the case where w is the j^{th} standard basis vector.

Figure 1 at the end of Section 2.3a gives added insight. To exploit symmetry one seeks changes of variable T so that the old problem $\mathcal{P}$ and new problem $\mathcal{Q}$ are *identical*: $\mathcal{P} = T^{-1}\mathcal{P}T$.

b) *Translation invariance*

If there are families of symmetries, one can obtain more information. We first discuss this for a linear differential equation with constant coefficients, $Lu = au'' + bu' + cu$. Here L commutes with *all* the *translation operators* T_α defined by $(T_\alpha u)(x) := u(x+\alpha)$. These translations T_α are a continuous group of symmetries: $T_\alpha T_\beta = T_{\alpha+\beta}$. The eigenfunctions of translations are just exponentials: $T_\alpha e^{cx} = \mu e^{cx}$, where $\mu = e^{c\alpha}$. We claim that these exponentials are also eigenfunctions of L. While this is simple to show directly, we prove more generally that this is true for any linear map L that commutes with all translations; some other instances are constant coefficient linear difference and linear partial differential equations (in this PDE case x, α, and c are vectors and cx becomes the inner product), and convolution equations.

Write $q(x;\lambda) := Le^{\lambda x}$. Since $T_\alpha e^{\lambda x} = e^{\lambda\alpha}e^{\lambda x}$, we have

$$T_\alpha Le^{\lambda x} = T_\alpha(q(x;\lambda)) = q(x+\alpha;\lambda) \qquad \text{and} \qquad LT_\alpha(e^{\lambda x}) = e^{\lambda\alpha}Le^{\lambda x} = e^{\lambda\alpha}q(x;\lambda).$$

Comparing these at $x = 0$, we see that if the linear map L commutes with translations, then $q(\alpha;\lambda) = q(0;\lambda)e^{\lambda\alpha}$ for any α. Equivalently, $q(x;\lambda) = q(0;\lambda)e^{\lambda x}$. Writing $Q(\lambda) := q(0;\lambda)$, we conclude

$$Le^{\lambda x} = Q(\lambda)e^{\lambda x}. \tag{35}$$

Thus $e^{\lambda x}$ is an eigenfunction of L for any λ, and the corresponding eigenvalue is $Q(\lambda)$.

Working formally, we apply (35) to find some solution of $Lu = f$. Write f and also seek a solution u as linear combinations of exponentials:

$$f(x) = \sum f_\lambda e^{\lambda x}, \qquad u(x) = \sum u_\lambda e^{\lambda x} \qquad \text{so} \qquad Lu = \sum u_\lambda Q(\lambda)e^{\lambda x}$$

(or integrate: $f(x) = \int f_\lambda e^{\lambda x}\,d\lambda$, etc.). To solve the homogeneous equation $Lu = 0$ use the roots of $Q(\lambda)$ while for the inhomogeneous equation use (35) and match coefficients to conclude that $u_\lambda = f_\lambda/Q(\lambda)$. Thus a solution is $u(x) = \sum\left[f_\lambda/Q(\lambda)\right]e^{\lambda x}$. One recognizes these formulas as the standard Fourier series/integrals and Laplace transform methods. This is why Fourier series and Fourier and Laplace transforms are so useful for constant coefficient differential equations. The value of $Q(\lambda)$ is determined separately for each problem. Since $Q(\lambda)$ appears in the denominator of the solution, its zeros play an important role, especially for partial differential operators, although we shall not pursue this further here. The point is that just by using translation invariance we know how to proceed.

As a quick application, return to the special case $Lu = au'' + bu' + cu$, where a, b and c are constants. Then $Le^{\lambda x} = (a\lambda^2 + b\lambda + c)e^{\lambda x}$, so $Q(\lambda) = a\lambda^2 + b\lambda + c$. In particular, if $Q(r) = 0$, then $u(x) = e^{rx}$ is obviously a solution of the homogeneous equation $Lu = 0$, while if $Q(r) \neq 0$, then $u(x) = e^{rx}/Q(r)$ is a particular solution of the inhomogeneous equation $Lu = e^{rx}$; if $Q(r) = 0$ but $Q'(r) \neq 0$, then one can take the derivative of (35) with respect to λ and evaluate at $\lambda = r$ to solve $Lu = e^{rx}$. Similarly, if r is a double root of $Q(\lambda) = 0$ then also $Q'(r) = 0$; here taking the derivative of equation (35) with respect to λ and evaluating at $\lambda = r$ reveals that $u(x) = xe^{\lambda x}$ is also a solution of the homogeneous equation, a fact that often is bewildering in elementary courses in differential equations.

We will look at this simple example a bit more. Let $S(t)$ be a fundamental matrix solution of the first order constant coefficient system $Lu := u' + Au = 0$, where A. This system is translation invariant so $S(t+\alpha)$ is also a solution for any α. Since the general solution has the form $S(t)C$ for some constant matrix C we know that $S(t+\alpha) = S(t)C$. Setting $t = 0$ gives $S(\alpha) = C$ so we deduce that the general exponential addition formula $S(t+\alpha) = S(t)S(\alpha)$ holds for more than the special case of $u' - u = 0$.[6] By writing $u'' + u = 0$ as a first order system, one finds that this general addition formula implies the usual formulas for $\sin(t+\alpha)$ and $\cos(t+\alpha)$. Further, since $S(t)S(-t) = I$, then $S^{-1}(t) = S(-t)$. Thus Green's function $G(t,\tau) = S(t)S^{-1}(\tau) = S(t-\tau)$.

There is an interesting cultural difference between the way mathematicians and physicists usually write the general solution of $u'' + u = 0$. Mathematicians write $u(x) = A\cos x + B\sin x$, which emphasizes the linearity of the space of solutions, while physicists write $u(x) = C\cos(x+\alpha)$, which emphasizes the translation invariance.

As an exercise apply translation invariance to develop the theory of second order linear difference equations with constant coefficients, $au_{n+2} + bu_{n+1} + cu_n = f(n)$. The Fibonacci [*c.* 1180–1250] sequence $u_{n+2} = u_{n+1} + u_n$, with initial conditions $u_0 = 0, u_1 = 1$, is a special case.

Invariance under multiplication $x \mapsto cx$ is related closely to translation invariance: if we let $x = e^z$, then translating z multiplies x by a constant. With this hint, one can treat the Euler differential operator $Lu = \alpha x^2 u'' + \beta x u' + \gamma u$, where α, β, γ are constants; this operator commutes with the stretching $x \mapsto cx$. Here the analog of the Fourier transform is called the Mellin transform.

The Laplace operator in Euclidean space is invariant under translations and orthogonal transformations; on a Riemannian manifold this property generalizes by the Laplacian being invariant under all isometries. The wave equation is invariant under Lorentz transformations (see the end of this Section). The basic point is that invariance under some large group automatically implies fundamental formulas and identities.

c) *More complicated group invariance*

In more complicated problems, there may be some symmetry but it may not be obvious to find or use. Sophus Lie [1842–99] created the theory of what we now call *Lie groups* to exploit symmetries to solve differential equations. His vision was to generalize Galois theory to differential equations. The resulting theory has been extraordinarily significant throughout mathematics. As our first example, observe that the differential equation

$$\frac{dy}{dx} = \frac{ax^2 + by^2}{cx^2 + dy^2} \qquad a, b, c, d \text{ constants}$$

is invariant if one makes the change of variable (a stretching) $x \mapsto \lambda x,\ y \mapsto \lambda y$ for any value of $\lambda > 0$. In other words, if $y = \varphi(x)$ is a solution, then so is $\lambda y = \varphi(\lambda x)$, that

[6] Conversely, if the square matrix $S(t)$ is differentiable and satisfies the functional equation $S(t+\alpha) = S(t)S(\alpha)$ for all α, then differentiating this with respect to t and setting $t = 0$ we conclude that S satisfies $S' + AS = 0$, where $A = -S'(0)$.

is $y=\varphi(\lambda x)/\lambda$. This motivates us to introduce a new variable that is invariant under this stretching: $w=y/x$. Then w satisfies $xw'=(a+bw^2)/(c+dw^2)-w$, which can be solved by separation of variables. The equation $dy/dx=(ax+by+p)/(cx+dy+q)$ has the symmetry of stretching from the point of intersection of the lines $ax+by+p=0$ and $cx+dy+q=0$. Lie showed that many complicated formulas one has for solving differential equations are but special instances of invariance under a family of symmetries. His work showed that a daunting bag of tricks that demoralize undergraduates were merely instances of exploiting symmetries. The next example is not as simple, so we'll be a bit more systematic.

Nonlinear equations of the form $\Delta u=f(x,u)$ arise frequently in applications. For instance the special cases where $f(x,u)$ has the forms $|x|^a u^b$ and $|x|^c e^u$ arise in astrophysics (Emden-Fowler equation), complex analysis, and conformal Riemannian geometry. We briefly discuss

$$\Delta u=|x|^c e^u \tag{36}$$

in $\mathbb{R}^n$ from the view of symmetry. While there are systematic approaches to seek symmetry, in practice one usually tries to guess; the method is of no help if finding symmetries is as difficult as solving the original problem.

For (36) the right side suggests we seek a symmetry group in the form $G:(x,u)\mapsto(\alpha x,\, u+\lambda)$, that is, we try the change of variables $\tilde{x}=\alpha x$, $\tilde{u}=u+\lambda$, where $\alpha>0$, λ are constants. Let $\tilde{\Delta}=\partial^2/\partial\tilde{x}_1^2+\cdots=\alpha^{-2}\Delta$ be the Laplacian in these new variables. Then $\tilde{u}(\tilde{x})$ is a solution of $\tilde{\Delta}\tilde{u}=[\alpha^{c+2}e^{\lambda}]^{-1}|\tilde{x}|^c e^{\tilde{u}}$. Thus if we pick $\alpha^{c+2}e^{\lambda}=1$, so $\lambda=-(c+2)\ln\alpha$, then $\tilde{u}(\tilde{x})$ is a solution of (36) for any value of α. In other words, if $u=\varphi(x)$ is a solution then so is $u(x)-(c+2)\ln\alpha=\varphi(\alpha x)$, that is, $u(x)=\varphi(\alpha x)+(c+2)\ln\alpha$ for any $\alpha>0$. The symmetry group is $G_\alpha:(x,u)\mapsto(\alpha x,u-(c+2)\ln\alpha)$. This is the identity map at $\alpha=1$.

To go further, recall that the Laplacian is invariant under the orthogonal group: if $u(x)$ is a solution, so is $u(Rx)$ for any orthogonal transformation R. It thus is reasonable to seek special solutions $u=u(r)$, where $r=|x|$, that are also invariant under the orthogonal group. Writing the Laplacian in spherical coordinates leads us to consider

$$u''+\tfrac{n-1}{r}u'=r^c e^u,$$

where $u'=du/dr$. We know this equation is invariant under the change of variables

$$\tilde{r}=\alpha r,\qquad \tilde{u}=u-(c+2)\ln\alpha. \tag{37}$$

For fixed r and u, as we vary α, (37) defines a curve in the $\tilde{r},\tilde{u}$ plane. It is natural to define new coordinates in which these curves are straight lines, say parallel to the vertical axis. We want one function $s = s(\tilde{r}(r,u,\alpha),\tilde{u}(r,u,\alpha)) = s(\alpha r, u-(c+2)\ln\alpha)$ that is constant on each of these curves; this function is used to select which of these curves one is on. The other function $v = v(\tilde{r}(r,u,\alpha),\tilde{u}(r,u,\alpha)) = v(\alpha r, u-(c+2)\ln\alpha)$ is used as a normalized parameter along these curves, chosen so that the directional derivative of v along these curves is one; see Figure 4. Thus, the conditions are

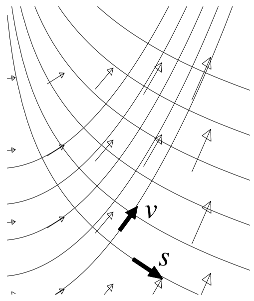

Figure 4

$$\text{(38)} \qquad \left.\frac{\partial s}{\partial\alpha}\right|_{\alpha=1} = 0 \qquad \text{and} \qquad \left.\frac{\partial v}{\partial\alpha}\right|_{\alpha=1} = 1.$$

By the chain rule these can be rewritten as

$$\text{(39)} \qquad rs_r-(c+2)s_u = 0 \qquad \text{and} \qquad rv_r-(c+2)v_u = 1,$$

where s_r etc. are the partial derivatives. Using the tangent vector field V to our curves,

$$V := \left.\frac{\partial\tilde{r}}{\partial\alpha}\right|_{\alpha=1}\frac{\partial}{\partial r} + \left.\frac{\partial\tilde{u}}{\partial\alpha}\right|_{\alpha=1}\frac{\partial}{\partial u} = r\frac{\partial}{\partial r} - (c+2)\frac{\partial}{\partial u},$$

we can rewrite (39) as

$$Vs = 0 \qquad \text{and} \qquad Vv = 1;$$

V is called the *infinitesimal generator* of the symmetry. In these new coordinates, by integrating (38)) the invariance (37) is simpler:

$$\text{(40)} \qquad \tilde{s} = s \qquad \text{and} \qquad \tilde{v} = v+\alpha.$$

An obvious particular solution of the second equation in (39) is $v = \ln r$; an equally obvious solution is $v = -u/(c+2)$, which would also work.

The first equation in (39) is straightforward to solve;[7] for variety we use an alternative approach to obtain $s(r,u)$. Eliminate α from the formulas (37) and find that $\tilde{u}+(c+2)\ln\tilde{r} = u+(c+2)\ln r$. Thus the function $s = (c+2)\ln r+u$ is constant along each of these curves. Since any function of s has the same property one can use this flexibility to choose a "simple" s. In these new coordinates, $s = u+(c+2)\ln r$, $v = \ln r$. After a computation that is not painless one finds that $v(s)$ satisfies

$$\ddot{v} = (n-2)[1-(c+2)\dot{v}]\dot{v}^2 - e^s\dot{v}^3,$$

where $\dot{v} = dv/ds$ and $\ddot{v} = d^2v/ds^2$. Since this does not involve v itself,[8] the substitution $w = \dot{v}$ gives a *first* order equation for $w(s)$, which simplifies significantly if $n = 2$, exactly the case of interest in applications.

[7] To solve $a(x,y)\psi_x+b(x,y)\psi_y = 0$ for $\psi(x,y)$, solve the ordinary differential equation $dy/dx = b/a$ and write its solution in the form $\psi(x,y) = C$, where C is the constant of integration. This $\psi(x,y)$ is a solution of the partial differential equation, as is any function of it. In our application the solution of $du/dr = -(c+2)/r$ is $u = -(c+2)\ln r+C$ so $\psi(r,u) = u+(c+2)\ln r$.

[8] Note that because of the invariance (40) in these variables, we knew in advance that this equations would not involve v.

It is a useful exercise to repeat this analysis for $\Delta u = |x|^a u^b$ in $\mathbb{R}^n$ and notice that the resulting equation simplifies dramatically when $(a+2)/(b-1) = (n-2)/4$, again exactly the situation of applications to physics and geometry. By using symmetry one can solve some problems that are otherwise impenetrable.

One impressive application of symmetry was G. I. Taylor's [1886–1975] computation of the energy in the first atomic explosion just by exploiting symmetry and taking measurements from publicly available photographs. For "security reasons" he did not have access to any technical data (see [B-K, Chapter 1] for an exposition). The monographs [B-K] and [Ol] show how to apply and exploit symmetry for ordinary and partial differential equations (it would be nice if there were a more accessible, less general, treatment).

Before the next example we should point out that in applications, invariance under the stretching $x \mapsto \lambda x$ arises frequently—since one uses stretchings to change to "dimensionless" variables (this is because the basic equations for any phenomena should be invariant if one changes from one set of units of measurement to another, say from "feet" to "meters"). Here is a small but useful mathematical application. For a bounded open set $\Omega \in \mathbb{R}^n$ say, generalizing the usual space C^1 (see also Section 3.2 below), for smooth functions u, which we assume have compact support in Ω, we define a similar norm using the L_p norm of the first derivatives:

$$\|u\|_{H_{1,p}(\Omega)} := \left[\int_\Omega |\nabla u(x)|^p \, dx\right]^{1/p}, \qquad p \geq 1,$$

and ask when the following inequality holds:

$$\sup_{z\in\Omega} |u(z)| \leq c(p,n,\Omega)\, \|u\|_{H_{1,p}(\Omega)}, \tag{41}$$

with the constant c independent of u (one should think of (41) as a version of the mean value theorem).

Since the left side of the inequality is invariant under stretching while, for most values of p the right side is not, we try a stretching to see what information it yields. For simplicity, say Ω contains the origin, so it contains some disk $\{|x| < a\} \in \mathbb{R}^n$, and let $\varphi(x)$ be a fixed smooth function that is zero for $|x| > a$ (but not identically zero). Then let $u(x) = \varphi(\lambda x)$ where $\lambda \geq 1$ is a constant. Computing both sides of (41) with this function we obtain

$$\sup_\Omega |\varphi| \leq c(p,n,\Omega)\, \lambda^{(p-n)/p}\, \|\varphi\|_{H_{1,p}(\Omega)}.$$

Since this is to hold for any $\lambda \geq 1$, we see that if $p - n < 0$, there is a contradiction if we let $\lambda \to \infty$. Thus, we conclude that $p \geq n$ is a necessary condition for inequality (41) to be valid. If $p \geq n$ then in fact the inequalities (41) do hold; they are called *Sobolev inequalities*. If $p = n$ this $H_{1,p}$ norm is invariant under stretchings, a fact that results in important and interesting properties.

d) *Noether's Theorem*

Most "natural" differential equations arise as Euler-Lagrange equations in the calculus of variations. Many believe one should always formulate fundamental equations using variational principles. E. Noether's [1882–1935] theorem shows how symmetry invariance of a variational problem implies basic identities, including conservation laws. While shorter direct proofs of these conservation laws might be found after one knows what to prove, there is a view that the symmetry is considerably deeper and more basic. Moreover, symmetry gives a way of finding new conservation laws.

To give a taste of the procedure we will deduce the standard "conservation of energy" for the vibrating string $\Omega = \{a < x < b\}$. A function $u(x,t)$ gives the displacement of a point $x \in \Omega$ at time t. The *wave equation* $u_{tt} = c^2 u_{xx}$, governs the motion; here c is the speed of sound. For simplicity we assume that $c = 1$. To eliminate the possibility of energy being added at the ends of the string, we will assume the string is fixed at the boundary, so $u(a,t) = u(b,t) = 0,\ t \geq 0$, as is typical for violin strings. In (15) we saw that the wave equation is the Euler-Lagrange equation for the functional

$$J[u] = \tfrac{1}{2}\int_\Omega \int_\alpha^\beta (u_t^2 - u_x^2)\,dx\,dt. \tag{42}$$

If we make the change of variables $\tilde{t} = t + \varepsilon$, since the integrand does not contain t explicitly, the functional J is invariant. Thus $dJ[u]/d\varepsilon\big|_{\varepsilon=0} = 0$. By an explicit computation we will show that this obvious fact implies conservation of energy.

With an eye toward generalization, it is useful to think of this as a change of variable in all the variables: $\tilde{t} = t + \varepsilon,\ \tilde{x} = x,\ \tilde{u} = u$ from (x,t,u) space to $(\tilde{x},\tilde{t},\tilde{u})$ space. This translation of t by ε takes the graph $u = u(x,t)$ into the graph $\tilde{u} = \tilde{u}(\tilde{x},\tilde{t};\varepsilon)$, thus $J[\tilde{u}(\tilde{x},\tilde{t};\varepsilon)] = J[u(x,t)]$. Because of this invariance, we clearly have $dJ[\tilde{u}]/d\varepsilon\big|_{\varepsilon=0} = 0$. Now

$$\begin{aligned} 0 = \left.\frac{dJ[\tilde{u}]}{d\varepsilon}\right|_{\varepsilon=0} &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \tfrac{1}{2}\int_\Omega \int_{\alpha+\varepsilon}^{\beta+\varepsilon} \left[u_{\tilde{t}}(x,\tilde{t}-\varepsilon)^2 - u_x(x,\tilde{t}-\varepsilon)^2\right]\,dx\,d\tilde{t} \\ &= \tfrac{1}{2}\int_\Omega \left[u_t(x,t)^2 - u_x(x,t)^2\right]\,dx\bigg|_{t=\alpha}^{t=\beta} + \int_\Omega\int_\alpha^\beta (-u_t u_{tt} + u_x u_{xt})\,dx\,dt. \end{aligned} \tag{43}$$

To go further, observe that in this last term $u_t u_{tt} = (u_t^2)_t - u_t u_{tt}$ and $u_x u_{xt} = (u_x u_t)_x - u_{xx} u_t$. Since we assumed $u(x,t)$ is an extremal of this functional, it satisfies the wave equation; thus the final integrand above is

$$-u_t u_{tt} + u_x u_{xt} = -(u_t^2)_t + (u_x u_t)_x + (u_{tt} - u_{xx})u_t = -(u_t^2)_t + (u_x u_t)_x.$$

We use this to simplify the last integral in (43) by evaluating the t integral in the first term and the x integral in the second:

(44)

$$\int_\Omega\int_\alpha^\beta [-u_t u_{tt} + u_x u_{xt}]\,dx\,dt = \int_\Omega\int_\alpha^\beta [-(u_t^2)_t + (u_x u_t)_x]\,dx\,dt = -\int_\Omega u_t(x,t)^2\,dx\bigg|_{t=\alpha}^{t=\beta} + 0$$

where in the last term we used that $u(x,t) = 0$ for x on the boundary of Ω (the ends of the string), so the velocity $u_t(x,t) = 0$ on the boundary of Ω.

Substituting (44)) in (43) we conclude that

$$0 = -\tfrac{1}{2}\int_\Omega \left[u_t(x,t)^2 + u_x(x,t)^2\right]\,dx\bigg|_{t=\alpha}^{t=\beta}.$$

Thus the function

$$E(t) := \frac{1}{2}\int_{\Omega}(u_t^2 + u_x^2)\,dx \equiv \text{constant} \tag{45}$$

is constant as a function of time. Since $E(t)$ is the energy, this formula is called "Conservation of Energy."

Similarly, for any functional of the form $J[u] = \frac{1}{2}\int_{\Omega}\int_{\alpha}^{\beta} F(x,u,u_t,u_x)\,dx\,dt$, where the integrand does not depend explicitly on t, identical reasoning gives $\int_{\Omega}(F_{u_t}u_t - F)\,dx =$ const. For more on Noether's Theorem see the references [G-F], [G-H], [B-K], and [Ol].

e) *Using symmetry for Pell's equation*

Here is another way to use symmetry. We want all the integer solutions of

$$x^2 - 2y^2 = 1. \tag{46}$$

By experimentation you quickly find the solution $x = 3,\ y = 2$. Are there any others? Can you find *all* the solutions? They are the integer lattice points on the hyperbola (46).

Writing $X := (x,y)$ and $Q(X) := x^2 - 2y^2$, seek a symmetry of the hyperbola $Q(X) = 1$ as a linear change of variables $R : (x,y) \mapsto (ax+by, cx+dy)$ defined by the matrix $R = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. We want R to have the property $Q(RX) = Q(X)$; in more formal language, we want the group of automorphisms R of the quadratic form Q. If we can find R, and if we have one solution $X_1 = (x_1, y_1)$ of $Q(X) = 1$, then $X_2 := RX_1 = (ax_1 + by_1,\ cx_1 + dy_1)$ is another solution since $Q(X_2) = Q(RX_1) = Q(X_1) = 1$. Thus, knowing R enables us to construct new solutions from old ones.

These automorphisms R embody the symmetries of the polynomial $Q(X)$, much as the rotations T (orthogonal transformations) embody the symmetries of the more familiar polynomial $P(X) := x^2 + y^2$ since $P(TX) = P(X)$. If X_1 is a point on a circle centered at the origin, then $X_2 := TX_1$ is another point on the same circle.

For our quadratic polynomial the obvious symmetries are $x \to \pm x$ and $y \to \pm y$. We want more. Since

$$Q(RX) = (ax+by)^2 - 2(cx+dy)^2 = (a^2 - 2c^2)x^2 + 2(ab - 2cd)xy + (b^2 - 2d^2)y^2,$$

the condition $Q(RX) = Q(X)$ means $a^2 - 2c^2 = 1,\ ab - 2cd = 0$, and $(b^2 - 2d^2) = -2$. If we pick a and c to satisfy the first of these, which is just the original equation (46), then the other two conditions imply $d = \pm a$ and $b = \pm 2c$. This yields all the symmetries R of our quadratic polynomial.

For our purposes it is enough to use the solution $(3,2)$ we found of (46) so $a = 3,\ c = 2,\ b = 4,\ d = 3$, and $R = \left(\begin{smallmatrix} 3 & 4 \\ 2 & 3 \end{smallmatrix}\right)$. We began with the solution $X_1 := (x_1, y_1) = (3,2)$. Using this we find the solutions $X_2 = RX_1 = (17,12),\ X_3 = RX_2 = (99,70)$, etc. of (46). Since $\det R = 1$ and the elements of R are integers, both the symmetry R and its inverse R^{-1} take integer lattice points to integer lattice points.

The mapping R has two basic geometric properties. To describe them take two points $V_1 := (x_1, y_1)$ and $V_2 := (x_2, y_2)$ both on the right $(x > 0)$ branch of the hyperbola $x^2 - 2y^2 = 1$. Call this right branch Γ, and say that V_1 is *below* V_2 (and write $V_1 \prec V_2$) if $y_1 < y_2$. The geometric properties are:

- *R preserves the branch*: if a point V is on Γ, then so is RV.
- *R preserves the order on* Γ: If $V_1 \prec V_2$ then $RV_1 \prec RV_2$.

Note that R^{-1} also has these properties. Since R is a continuous map from the hyperbola to itself, by connectedness, it maps the right branch, Γ, either to itself or to the left branch. Checking the image of one point, say $(1,0)$ we see that the image is in Γ. Moreover, since R is invertible as a map of the whole plane, its restriction to Γ is invertible. Therefore it is either monotonic increasing or decreasing as a function of the y coordinate on Γ. Again checking the image of $(1,0)$, we conclude that the restriction of R to Γ is an increasing function of the y coordinate. This implies that R preserves the order on Γ.

Our particular solution $X_1 := (3,2)$ is the positive integral solution with the *smallest* possible positive value for y_1. Writing $X_0 = (1,0)$, this means $X_0 \prec X_1$ and there is no other integral solution between X_0 and X_1. Since $Q(RX_1) = Q(X_1) = 1$ we see that $X_2 := RX_1 = (17,12)$ is also a solution of (46). Similarly $X_k := (x_k, y_k) = RX_{k-1} = R^k X_0$ are all positive integer solutions for any positive integer k. These solutions are distinct since their y coordinates are increasing, so $X_k \prec X_{k+1}$.

Moreover, these are *all* the positive integral solutions. If there were another, Z, then for some k we have $X_k \prec Z \prec X_{k+1}$. Therefore $R^{-1}Z$ is yet another solution and because R preserves the order of the points on the hyperbola,

$$X_{k-1} = R^{-1}X_k \prec R^{-1}Z \prec R^{-1}X_{k+1} = X_k.$$

Continuing, we obtain a solution $R^{-k}Z$ between X_0 and X_1 since

$$X_0 = R^{-k}X_k \prec R^{-k}Z \prec R^{-k}X_{k+1} = X_1.$$

This contradicts the fact that $X_1 = (x_1, y_1) = (3,2)$ was the positive solution whose second coordinate was as small as possible. We conclude that $X_k = R^k X_0$, that is, the *orbit* of X_0 after repeated action by R, are all of the integer solutions.

The matrix R^k can be computed explicitly by first diagonalizing it. This gives $R^k = S\Lambda^k S^{-1}$, where Λ is the diagonal matrix of eigenvalues $3 \pm 2\sqrt{2}$ of R and S is the matrix whose columns are the corresponding eigenvectors $(\pm\sqrt{2}, 1)$; these vectors also determine the asymptotes of the hyperbola. Thus $X_k = R^k X_0$ has the formula

$$X_k = \left(\frac{(3+2\sqrt{2})^k + (3-2\sqrt{2})^k}{2}, \frac{(3+2\sqrt{2})^k - (3-2\sqrt{2})^k}{2\sqrt{2}} \right), \tag{47}$$

which shows that explicit formulas may be more complicated—and maybe less desirable—than you might anticipate. Perhaps of greater value, this formula leads us to define R^t, for $-\infty < t < \infty$, by the rule $R^t = S\Lambda^t S^{-1}$, so $R^{s+t} = R^s R^t$. If we let $X(t) = R^t X_0$, and write $X(t) = (x(t), y(t))$, then from (47) with k replaced by t we see that $x(t) = \frac{1}{2}(\alpha^t + \alpha^{-t})$ and, $y(t) = \frac{1}{2\sqrt{2}}(\alpha^t - \alpha^{-t})$, where $\alpha = 3 + 2\sqrt{2}$. By a straightforward computation one can verify that $x(t)^2 - 2y(t)^2 = 1$, that is, the points $X(t)$ are all on our hyperbola. It is now evident that $x(t) \geq 1$ and $dy/dt > 0$ so the "orbit" of $X(t)$ is the entire right branch Γ of the hyperbola with $y(t)$ an increasing function of t. Thus $X(s) \prec X(t)$ if and only if $s < t$. As a bonus, we see that *every* symmetry of the right branch of the hyperbola $x^2 - 2y^2 = 1$ has the form R^t for some real t.

One can use this to find all integer solutions of $x^2 - 2y^2 = k$ for integers k: assuming one has some solution one gets all solutions. Moreover this works for all "Pell" equations: $x^2 - Dy^2 = k$ with $D > 0$ not a perfect square. For a given k, once one finds some particular solution (x,y) all the others can be found using the solutions of $x^2-Dy^2=1$. For our example we found the particular solution by trial and error; in general there may *not* be any solution; for instance, there is no solution of $x^2-2y^2=3$ since there is no non-trivial solution in the integers mod3. One constructive approach that always works for the special case $x^2-Dy^2=1$ uses continued fractions (see [Da], [N-Z-M]), another (non-constructive) uses Minkowski's [1864–1909] geometry of numbers (see [Art]).

An essentially identical computation to finding the symmetries of x^2-2y^2 yields all linear changes of variable $x'=\alpha x+\beta t$, $t'=\gamma x+\delta t$ that preserve the wave operator $\partial^2/\partial t^2-c^2\partial^2/\partial x^2$, where c is a constant (the speed of sound or light). By the chain rule,

$$u_{tt}-c^2u_{xx}=(\delta^2-c^2\gamma^2)u_{t't'}+2(\beta\delta-c^2\alpha\gamma)u_{x't'}+(\beta^2-c^2\alpha^2)u_{x'x'}.$$

Thus we want $\delta^2-c^2\gamma^2=1$, $\beta\delta-c^2\alpha\gamma=0$, and $\beta^2-c^2\alpha^2=-c^2$. First pick γ and δ so that $\delta^2-c^2\gamma^2=1$, and then let $\beta=\pm c^2\gamma$, $\alpha=\pm\delta$. To preserve orientation we use the $+$ signs. Since $c^2\alpha^2-\beta^2=c^2$ and $\cosh^2\sigma-\sinh^2\sigma=1$, it is traditional to write $\alpha=\cosh\sigma$, $\beta=c\sinh\sigma$. For any real σ the transformation

$$\begin{aligned} x'&=(\cosh\sigma)x+(c\sinh\sigma)t\\ t'&=(\frac{1}{c}\sinh\sigma)x+(\cosh\sigma)t \end{aligned} \tag{48}$$

preserves the wave operator. This is called a *Lorentz transformation*. Lorentz [1853–1928] transformations also preserve arc length $ds^2:=dx'^2-c^2\,dt'^2=dx^2-c^2\,dt^2$ in space-time and are fundamental in the study of the wave operator and special relativity.

In special relativity it is enlightening to replace the parameter σ in (48) by one that is physically more meaningful. If the x-axis moves with constant velocity V relative to the x'-axis, for an observer on the x'-axis, $x'/t'=V$ is the constant velocity of the origin $x=0$ of the x-axis. But from (48) with $x=0$

$$V=\frac{x'}{t'}=c\tanh\sigma,$$

so $\sinh\sigma=(V/c)/\sqrt{1-(V/c)^2}$ and $\cosh\sigma=1/\sqrt{1-(V/c)^2}$. We can use this to rewrite the Lorentz transformation (48) in terms of the velocity V as

$$x'=\frac{x+Vt}{\sqrt{1-(V/c)^2}}\qquad t'=\frac{(V/c^2)x+t}{\sqrt{1-(V/c)^2}}.$$

It is physically obvious that to get the inverse transformation just replace V by $-V$.

3 Some Procedures To Prove Existence

Existence of a solution of an equation may be approached in different ways. One should first try to find a "simple" expression for the solution, perhaps using some of the procedures discussed already. The following discussion assumes this has been used as much as possible.

There are two types of existence procedures: those that construct a specific solution, and those that merely prove a solution exists. As examples, I present one constructive approach and two purely existential approaches to proving the existence of a solution. Recall

Hermann Weyl's [1885–1955]: "Whenever you can settle a question by explicit construction, be not satisfied with purely existential arguments." In the light of this dictum it is useful to reflect on the constructive and non-constructive approaches discussed in Section 2.2 for solving $ax \equiv b \pmod{m}$.

3.1 **Iteration methods**

A frequent procedure is to begin with a simpler problem that one knows how to solve and use that to solve nearby more complicated problems. Physicists and engineers call this "perturbation theory." Within mathematics the standard examples of these are iterative proofs of the implicit and inverse function theorems, and the existence of a solution of an ordinary differential equation. Often mathematicians refer to this method as finding a fixed point of a "contracting map"—but when one examines the proof, the essence is a simple iteration procedure (see [K-F]).

Although iterative methods were developed primarily for nonlinear problems, they can be important even in finite dimensional linear algebra. Here is an example. Say you know the inverse of a matrix A and someone gives you a matrix B that is almost the same as A. One suspects that B^{-1} will be near A^{-1}. This situation arises in models of the economy where the matrix $A = (a_{ij})$ may be very large, say with 10,000 rows and columns. Perhaps one identifies the 10,000 most significant ingredients in the economy, say steel, oil, wheat, electricity, cotton, the average hourly wage of a worker, etc. Then a_{ij} may represent the effect of increasing the cost of the i^{th} ingredient on the cost of the j^{th} ingredient. For instance, if one increases the cost of oil by \$1 per barrel, this will increase the cost of steel a certain amount. The matrix B may be the version A obtained from the next month's data.

This linear algebra problem is so large that it is best treated using analysis. The first step is to use the idea in Section 2.3: find a simpler equivalent problem. Write

$$B = A - (A - B) = A[I - A^{-1}(A - B)] = A(I - C),$$

where $C = A^{-1}(A - B)$ is presumably small since we assumed that B is near A. Then $B^{-1} = [I - C]^{-1}A^{-1}$, so all we need to do is compute the inverse of $I - C$, that is, we want a matrix D so that $(I - C)D = I$. Thus we have reduced to the special case when A is the identity matrix and $B = I - C$. Since C is small, we rewrite $(I - C)D = I$ as $D = I + CD$ and use the successive approximations $D_{k+1} = I + CD_k$, with the initial guess $D_0 = I$. This gives

$$D_1 = I + C, \qquad D_2 = I + C + C^2, \qquad D_3 = I + C + C^2 + C^3, \quad \text{etc.}$$

If C is small, then by picking k large D_k is an approximation to $(I - C)^{-1}$. This should not surprise us since we know the Taylor series for $1/(1 - x)$ for small x.

In computational problems, one may be able to use a different iteration method that converges faster. Newton's method is an example. For instance, with the usual method taught in schools for finding square roots (really just a version of preceding iteration method), you get one additional decimal place at each iteration, while with Newton's method you get *double* the number of decimal places with each iteration (see [D-B, Sec. 6.3]).

3.2 **Variational methods**

An example illustrates the issues vividly. Say we want to solve the system of equations

$$\begin{aligned} x^3 + 2xy - 3y\cos x e^{\sin x} &= -7 \\ y^5 + x^2 - 3\,e^{\sin x} &= \quad 5 \end{aligned}$$

Is there a solution? Without further insight this may not be obvious. But these two equations state that the gradient of the function

$$u(x,y) := \tfrac{1}{4}x^4 + \tfrac{1}{6}y^6 + x^2y - 3y\,e^{\sin x} + 7x - 5y$$

is zero. Thus, the solutions of our equation correspond to the critical points of $u(x,y)$. It is obvious that as one goes far from the origin then u becomes large. Thus, there is some point (x_0,y_0) where u takes on its minimum value. This minimum gives one solution of our equations. To determine if there are others would require a more detailed investigation.

This approach is a useful technique for proving that certain differential equations always have at least one solution. The method is called the "direct method in the calculus of variations." By the method of Section 2.2 a critical point of the functional

$$J(u) = \tfrac{1}{2}\iint_{\Omega}(u_x^2 + u_y^2)\,dx\,dy. \tag{49}$$

with $u = f$ on the boundary of Ω is a solution of the Laplace Equation $\Delta u = 0$ in a region Ω.

Following the example at the beginning of this section, to find a solution of $\Delta u = 0$, we can seek a minimum u of J. Since the functional J is non-negative, this leads one to assert that it attains its minimum at some function u, and proves the existence of a solution of $\Delta u = 0$ with the prescribed boundary values. This assertion is called *Dirichlet's Principle*.

After Riemann [1826–66] dramatically applied this reasoning in his work on complex analysis, Weierstrass [1815–96] pointed out this "principle" is *false* since he exhibited a similar functional J that only has an infimum and does *not* attain a minimum value in the class of admissible functions. Nonetheless, everyone—including Weierstrass—believed that Riemann's results were essentially correct. This forced mathematicians to develop the concept of compactness in function spaces, where is it considerably more subtle than in Euclidean space. The gap remained until Hilbert's work in 1901 and 1909. In this context, it is interesting to note Nietzsche's remark: "Great men's errors are to venerated as more fruitful than little men's truths".

3.3 Fixed point methods

Another example. Say you want to solve the system of equations

$$\begin{aligned} 3x - \;\; 5y &= \frac{2x + ye^{2-\sin xy}}{7 + x^2 + y^4} - 13 \\ 2x + 71y &= 9 - \cos(xy + 19e^{x-5y}) \end{aligned}$$

Is there at least one solution? Again, to most people this is not immediately obvious. You look at the equations . . . The equations look at you.

Eventually you may be led to write this in the form $LX = F(X)$, where $X = (x,y)$, L is the 2×2 matrix on the left side, and $F(X)$ is the nonlinear right side. The key observation is that the vector function $F(X)$ is bounded independently of X. In fact $\|F(X)\| \le 100$

(the size of the bound is unimportant for us). Moreover, the matrix L is invertible, so we can rewrite our equations in the symbolic form

$$X = T(X) \quad \text{where} \quad T(X) = L^{-1}F(X).$$

If we view $T(X)$ as a map from the plane $\mathbb{R}^2$ to itself, then the equation $X = T(X)$ means that the solution X we seek is a *fixed point* of the map T. Since $\|F(X)\| \leq 100$, we know that $\|T(X)\| \leq R$ for some constant R that is independent of X (we can let $R = 10{,}000$, but that is irrelevant for our immediate concerns). Thus we have found the *a priori* inequality: if a solution of our equation exists, it must lie in the closed disk $B = \{\|X\| \leq R\}$. Since T maps any point X into B, in particular it maps B into B.

Now we can invoke the Brouwer [1881–1966] fixed point theorem, a result customarily proved in topology courses (see [doC, p. 75] for a slick proof using Stokes' theorem). It asserts that any continuous map of a closed disk to itself must have at least one fixed point. This fixed point is the solution we seek.

The Schauder fixed point theorem generalizes the Brouwer theorem to infinite dimensional spaces. This generalization requires an additional compactness assumption. If B is a Banach space and $S \subset B$, then a continuous map $T : S \mapsto B$ is *compact* if for any bounded set $Q \subset S$ the closed set $\overline{f(Q)}$ is compact. For example, consider the Banach spaces $C(S^1)$ and $C^1(S^1)$ of 2π-periodic continuous functions and periodic continuously differentiable functions on the circle S^1 with the usual norms

$$\|u\|_{C(S^1)} = \max_{0 \leq x \leq 2\pi} |u(x)| \quad \text{and} \quad \|u\|_{C^1(S^1)} = \max_{0 \leq x \leq 2\pi} |u(x)| + \max_{0 \leq x \leq 2\pi} |u'(x)|.$$

We should (but will not) write C_{periodic} to emphasize the periodicity. The Arzelá-Ascoli theorem implies that the identity map $id : C^1(S^1) \hookrightarrow C(S^1)$ is compact. The Schauder fixed point theorem says that if $S \subset B$ is a closed, convex, bounded set and if $T : S \to S$ is a compact map, then T has a fixed point (see [13, p. 32]). Schauder devised it specifically for partial differential operators. As an application we prove the existence of at least one periodic solution $u(x)$ with period 2π of

$$u' + u = F(x, u),$$

assuming only that $F(x, s)$ is a smooth function, periodic with period 2π in x and uniformly bounded, $|F(x, s)| \leq k$, where the constant k is independent of x and s.

A key observation is that the linear equation $Lu = u' + u = f(x)$ has a unique 2π periodic solution for any smooth periodic function $f(x)$. A direct computation gives

$$u(x) = \frac{1}{e^{2\pi} - 1} \int_0^{2\pi} e^t f(t)\, dt + \int_0^x e^{t-x} f(t)\, dt$$

(solve for $u(x)$ as usual—see (13)—and then pick the constant of integration, $u(0)$, to force the periodicity: $u(2\pi) = u(0)$). This formula also yields the inequality $\|u\|_{C(S^1)} \leq \|f\|_{C(S^1)} = \|Lu\|_{C(S^1)}$. However $|u'| = |Lu - u| \leq |Lu| + |u|$ so we obtain the estimate

$$\|u\|_{C^1(S^1)} \leq 3\|Lu\|_{C(S^1)}. \tag{50}$$

This asserts that $L^{-1} : C(S^1) \to C^1(S^1)$ is a continuous map. Rewrite our problem as $u = L^{-1}F(x, u)$. Thus we seek a fixed point of the map $T(u) := L^{-1}F(x, u)$. Since we defined T as the composition

$$C(S^1) \xrightarrow{F} C(S^1) \xrightarrow{L^{-1}} C^1(S^1) \xhookrightarrow{id} C(S^1),$$

it is a compact map. Because $F(x,s)$ is bounded, then for some constant K,

$$\|T(u)\|_{C(S^1)} \leq K \quad \text{for all} \quad u \in C(S^1).$$

This proves *a priori* that any solution u of this problem must satisfy

$$\|u\|_{C(S^1)} = \|T(u)\|_{C(S^1)} \leq K.$$

Thus let B be the ball

$$B := \{u \in C(S^1) : \|u\|_{C(S^1)} \leq K\}.$$

The Schauder theorem shows there is at least one periodic solution $u \in B$ and $u \in C^1(S^1)$. Using a bootstrap argument, if $F(x,s)$ is smooth, then so is this solution u.

There is a similar result for $Lu := -\Delta u + cu = F(x,u)$ with various boundary conditions, assuming L is invertible and F is bounded. However one must use more complicated function spaces, such as Sobolev spaces, to prove an analogue of the fundamental inequality (50).

4 An Open Question

One is not surprised to see a seemingly elementary unsolved problem in number theory. It is less well-known that there are many interesting and simple-looking nonlinear partial differential equation about which little is known. Let $f(x,y)$ be a smooth function. Is there always at least one solution $u(x,y)$ of the Monge-Ampère equation (Monge [1746-1818], Ampère [1775–1836])

$$u_{xx}u_{yy} - u_{xy}^2 = f(x,y)? \tag{51}$$

This is a modest question. We seek some solution in a possibly small neighborhood of the origin; no additional conditions such as initial or boundary conditions are imposed. Yet we still do not know the answer. Many cases have been treated. If $f(x,y)$ has a power series expansion, we can invoke the Cauchy-Kowalewskaya theorem to get a power series solution. If $f(0,0) > 0$, we can use the theory of elliptic partial differential equations to prove that a solution exists, while if $f(0,0) < 0$ we appeal to the theory of hyperbolic equations. The difficult case is when $f(0,0) = 0$. This case has also been treated if either $f(x,y) \geq 0$ near the origin, or if $\nabla f(0,0) \neq 0$, [Lin1], [Lin2]. Nothing more is known. Perhaps there are smooth functions with $f(0,0) = 0$ for which no solutions exist.

A similar differential equation arises in geometry. Locally, an abstract two dimensional surface with a Riemannian metric is a neighborhood of the origin in the u,v plane where one specifies the element of arc length

$$ds^2 = E(u,v)\,du^2 + 2F(u,v)\,du\,dv + G(u,v)\,dv^2 \tag{52}$$

of curves in that neighborhood. You always get an arc length of this form if you consider the curves $u(t)$, $v(t)$ on a two-dimensional surface with local coordinates u,v in $\mathbb{R}^n$. Does this give all possible abstract Riemannian metrics for the special case of surfaces in $\mathbb{R}^3$? In other words, given any arc length ds^2 of the form (52), locally can one always find a surface $x = x(u,v)$, $y = y(u,v)$, $z = z(u,v)$ in $\mathbb{R}^3$ having this as its arc length? More briefly, can every abstract two-dimensional Riemannian manifold be locally isometrically embedded in $\mathbb{R}^3$? One can show that there is a surface in $\mathbb{R}^4$ having this arc length, but the more interesting $\mathbb{R}^3$ case is still open. In one approach, the partial differential equation to be solved is essentially (51). Here the Gauss curvature $K(x,y)$ plays the role of the function $f(x,y)$, so we know there is a local embedding if $K(0,0) \neq 0$. The difficult case remaining is when $K(0,0) = 0$.

Problems such as this are challenges for the future.

I find that the harder I work, the more luck I seem to have.
Thomas Jefferson (1743-1826)

When I am working on a problem I never think about beauty. I only think about how to solve the problem. But when I have finished, if the solution is not beautiful, I know it is wrong.
Buckminster Fuller (1895–1983)

Some References

[A-K-L] Aleksandrov, A. D., Kolmogorov, A. N., Lavrent'ev, *Mathematics, Its Content, Methods, and Meaning*, 3 Vols, M.I.T. Press, Cambridge, Mass (U.S.A.), 1963.

[Arn] Arnold, V.I., *The Theory of Singularities and its Applications*, Lezioni Fermiane, Accademia Nazionale del Lincei Scuola Normale Superiore, Pisa, published by the Press of the University Cambridge, England 1991

[Art] Artin, Michael, *Algebra*, Prentice-Hall, New Jersey, 1991.

[Be] Bellman, Richard, *Introduction to Matrix Analysis*, McGraw-Hill, 1960, reprinted (1995) by SIAM in their series "Classics in Applied Mathematics."

[B-K] Bluman, George W., and Kumei, Sukeyuki, *Symmetries and Differential Equations*, Springer-Verlag, 1989, New York. Applied Mathematical Sciences **81**.

[C-H] Castrigiano, Domenico P.L. and Hayes, Sandra A., *Catastrophe theory*, Addison-Wesley Pub. Co., Advanced Book Program, 1993, Reading, Mass.

[D-B] Dahlquist, G., and Björck, Å., *Numerical Methods*, Prentice-Hall Series in Automatic Computation, Prentice-Hall, 1974.

[Da] Davenport, H., *The Higher Arithmetic*, Harper & Row, New York, 1952, 6th edition published by Cambridge Univ. Press, 1992.

[doC] do Carmo, Manfredo P., *Differential forms and applications*, Universitext, Springer-Verlag, New York, 1994.

[Do] Donaldson, S.K., “The Seiberg-Witten equations and 4-manifold topology”, *Bull. AMS*, **33** (1996), 45-70.

[Eb] Ebbinghaus, H.D., et al. *Numbers*, Springer-Verlag, 1991, New York. Graduate texts in mathematics: Readings in mathematics **123**. Translated from *Zahlen, Grundwissen Mathematik*. [A series of articles (by real mathematicians) tracing the development of “numbers” from integers to quaternions and beyond. It should be better known.]

[F] Forsythe, George E., “Pitfalls in Computation, or Why a Math Book Isn’t Enough”, *Amer. Math. M.*, **77** (1970), pp. 931–956.

[F-M-M] Forsythe, George E., Malcom, Michael A., Moler, Cleve B., *Computer Methods for Mathematical Computations*, Prentice Hall, 1977.

[Ful] Fulton,William *Algebraic Curves*, Benjamin (Addison-Wesley), 1969.

[G-F] Gelfand, I.M., and Fomin, S.V., *Calculus of Variations*, Prentice Hall, 1963 (translated from the Russian).

[G-H] Giaquinta, M. and Hildebrandt, S., *Calculus of Variations*, Grundlehren Vols. 310 and 311, Springer-Verlag, 1996, Berlin

[G-S] Golubitsky, Martin, and Schaeffer, David G. *Singularities and groups in bifurcation theory*, Applied mathematical sciences ; 51, 69, Springer-Verlag, 1985, New York.

[H-K] Hale, Jack and Koçak, Hüseyin, *Dynamics and Bifurcations*, Texts in Applied Mathematics **3**, Springer-Verlag, 1991

[H-T] Hildebrandt, S., and Tromba, A., *The Parsimonious Universe*, Springer-Verlag, 1996, New York (an earlier version: *Mathematics and Optimal Form* was published in the Scientific American Library, 1985).

[K] Kazdan, Jerry L.,*Prescribing the Curvature of a Riemannian Manifold*, CBMS **57**, American Math. Soc., 1985.

[K-F] Kolmogoroff, Andrei Nikolaevich, and Fomin, Sergei Vasil’evich, *Introductory real analysis*, Dover Publications, New York, N.Y., 1975, also published under the title *Elements of the theory of functions and functional analysis* by Graylock Press, Rochester, N.Y., 1957

[Lin1] Lin, Chang Shou, “The local isometric embedding in $\mathbf{R}^3$ of 2-dimensional Riemannian manifolds with nonnegative curvature,”*J. Diff. Geom.* **21** (1985), 213–230.

[Lin2] __________, “The local isometric embedding in $\mathbf{R}^3$ of two-dimensional Riemannian manifolds with Gaussian curvature changing sign cleanly,” *Comm. Pure Appl. Math.,* **39** (1986), 867–887.

[McO] McOwen, Robert, *Partial Differential Equations*, Prentice-Hall, 1995.

[N] Nirenberg, L., *Topics in nonlinear functional analysis*, New York University Lecture Notes, 1974.

[N-Z-M] Niven, I., Zuckerman, H., and Montgomery, H., *An Introduction to the Theory of Numbers*, Fifth Edition, John Wiley, 1991, New York.

[Ol] Olver, Peter J., *Applications of Lie groups to differential equations*, 2nd ed., in the series Graduate texts in mathematics, **107**, Springer-Verlag, 1993, New York.

[P] Polya, George, *How To Solve It*, 2nd. ed., Princeton University Press, Princeton, N.J., 1973. A classic. [While its scope is broader than just solving equations, its advice is valuable.]

[R-R] Renardy, Michael, and Rogers, Robert C., *An Introduction to Partial Differential Equations*, Springer Texts in Applied Mathematics **13**, Springer-Verlag, New York, 1993.

[Wa] Walker, Robert J., *Algebraic Curves*, Princeton Univ. Press, 1950, reprinted by Dover, 1962.

[War] Warner, Frank W., *Foundations of Differentiable Manifolds and Lie Groups*, Scott, Foresman & Co., 1971. Second edition published by Springer-Verlag in their series Graduate Texts in Mathematics, Vol. 94.

[Wf] Wilf, Herbert S., *Mathematics for the Physical Sciences*, Dover Publications, 1978 (reprinted from an earlier edition).

[Wi] Wilie, C. Ray, and Barrett, Louis, *Advanced Engineering Mathematics*, 6^{th} ed, McGraw-Hill, New York, 1995.

Jerry L. Kazdan
Department of Mathematics
University of Pennsylvania
Philadelphia, PA 19104-6395
email: kazdan@math.upenn.edu