PROOF PROJECTS FOR TEACHERS

FRANK QUINN

Abstract. This essay illustrates the contemporary mathematical approach to concept formation and abstraction, using topics important in elementary education. The October draft is incomplete and will be significantly revised.

Contents

Date: Draft, October 7, 2009.

Introduction

This note[1] outlines projects for college students who may become elementary or secondary teachers. It was written to test and illustrate ideas in [5] about proofs, definitions, abstractions and mathematical methods. Accordingly, it is a resource or starting point, and not intended to be used in this form.

Each project should be done as a unit without interruption. If students are immersed in a topic the ideas will become familiar and easy to work with. If there are interruptions then it is harder to develop this familiarity and the work will be harder. Further, in the sections that are covered all problems should be worked. Skipping material also slows development of familiarity and will make later work harder.

Problems could be worked in groups. It is a good ideas to go over proofs in groups to be sure that the sense is communicated correctly, see §1.4.1 (Style in Short–Form Proofs).

Fractions in commutative rings are developed in §1. The general treatment includes polynomial fractions (rational functions) and many other things for little more effort than needed for a careful development of integer fractions. The section on Grothendieck groups, §1.12, provides interesting prospective: the same construction with additive rather than multiplicative notation, and what happens when you allow division by zero.

A formula for the area of the region enclosed by a closed piecewise linear path is developed and explored in §2. The development is relatively elaborate, in part because students do not have a previous definition to compare it with. Winding numbers are used to describe the general case as "area counted with multiplicity", and the polyhedral Jordan Curve theorem comes out of the development. Finally sections §3.1.1, 3.1.2 explore possible extensions of the formula. They provide prospective on both the area formula and the way mathematics is done.

Comments on structure and learning of mathematics directed to students are in footnotes or the text. Comments for educators are in §5, and mainly concern practical aspects of learning programs that include formal definitions and proofs.

1. Fractions

We investigate fractions in a general context that includes both integer fractions and rational functions (fractions of polynomials).

1.1. **Commutative Rings.** We suppose R is a commutative ring. This means there are addition and multiplication operations that obey all the same basic rules as the operations in the integers. There are, of course, axiomatic formulations of these rules (commutative, distributive etc.) but they are already familiar so you can work without thinking about them explicitly[2]. Examples are given in §1.2.

[1] Written for Fou-Lai Lin, with thanks for hospitality at ICMI Study 19, May 2009.
[2] See comment in §1.5

1.1.1. *Notation Comment.* A *ring* (without "commutative") has addition and multiplication operations that obey the basic rules *except* multiplication may not commute (ab is generally different from ba). After long experience mathematicians have found that this is a more basic structure so it gets the short name. If multiplication does commute then "commutative" is added.

1.2. **Examples of Commutative Rings.**

(1) the Integers, $\mathbf{Z}$
(2) the Integers modulo a number n, denoted $\mathbf{Z}/n$
(3) polynomials with real coefficients,

$$\mathbf{R}[x] = \{\sum_{i=0}^{n} a_i x_i \mid \text{some finite } n \text{ and } a_i \text{ real}\}.$$

(4) Laurent polynomials (negative exponents are allowed),

$$\mathbf{R}[x, x^{-1}] = \{\sum_{i=-n}^{n} a_i x_i \mid \text{some finite } n \text{ and } a_i \text{ real}\}.$$

(5) formal power series ("infinite polynomials") denoted

$$\mathbf{R}[[x]] = \{\sum_{i=0}^{\infty} a_i x_i \mid a_i \text{ real}\}.$$

The last three are examples of general constructions: if one starts with a commutative ring R then polynomials, series etc. with coefficients in R give new commutative rings denoted $R[x]$, etc. Real–coefficient polynomials are given as the example because these are important in the study of functions and calculus but much of what we do holds more generally.

Because we are working abstractly we can study fractions in all these examples at once with the same effort needed to study integer fractions.

1.3. **Preliminary Definition of Fractions.**

The key property of a fraction is that $b\frac{a}{b} = a$. We make this official:

Definition, preliminary version If a and b are in R then $\frac{a}{b}$ is a name for the solution of the equation $b\phi = a$.

This is "preliminary" because there are serious problems with it, see §1.7, and the final version has a restriction on b to ensure it makes sense. However this discussion is postponed until after some practice work with inverses.

1.3.1. *Notation Comment.* The definition says $\frac{a}{b}$ is a *name* for an object. Objects can have several names. For instance the integer fraction $\frac{1}{4}$ also has the decimal name 0.25. The fraction name encodes the equation it satisfies, just as $\sqrt{a}$ encodes the fact that it is the nonnegative solution of the equation $x^2 = a$.

If we care about the connection to the integers then $\frac{1}{4}$ is a good name. For example $\frac{1}{4} + \frac{7}{28} + \frac{197}{23}$ is a name for a number. It has a fraction name and we would want this if we care about the connection to the integers. If we do not care about this connection then it would be easier to use decimal names.

1.4. **Inverses.** If the equation $bx = 1$ has a solution *inside* R then it is called the *inverse* and written b^{-1}.

For example, every nonzero real or rational number has an inverse in the same ring. Examples and significance of inverses is discussed after the following problem.

Sample Problem: Fractions and Inverses: Show that if b has an inverse then $\frac{a}{b} = a\,b^{-1}$.

To illustrate what "show" means we give the solution.
Solution, Long form:

(1) The hypothesis is that b has an inverse, which unpacks to: there is b^{-1} with $b^{-1}b = 1$.
(2) The conclusion unpacks to: $\phi = a\,b^{-1}$ is a solution for the equation $b\phi = a$.
(3) We check to see if the unpacked version of the conclusion is true for $\phi = b^{-1}$: $b(a\,b^{-1}) = (bb^{-1})a = (1)a = a$. The information in (1) is used in the middle step. Since the unpacked version of the conclusion is true so is the packed version, and we are finished[3].

"Unpacking" is described below. Note the arithmetic in the last step is business as usual, even though it is going on in some unspecified commutative ring. We do *not* spell out the axioms needed to justify it. Also, the first two steps are routine unpackings that we know in advance that we will have to do. They require a bit of care but no real thought. After that the core step (3) is easy.

Solution, Short form: According to the definition of $\frac{a}{b}$ we need to show $b(a\,b^{-1}) = a$. But this is immediate from the definition of the inverse and standard rules of arithmetic.

1.4.1. *Style in Short–Form Proofs.* Students should carefully compare the long and short forms of the proof above. The long form is the "official" version while the short form is a compressed version that can be routinely expanded to get the long form. However short–form proofs are acceptable only if they really can be expanded. This means *style* is important in short–form proofs: minor errors in calculation or use of words may cause doubt that the person who wrote the proof really could have written a valid long–form proof. Students who have trouble with style in short–form proofs should practice writing out long forms and then compress them. After some practice they should be able to write directly in the short form.

1.5. **Unpacking Definitions.** "Unpacking" is the use of definitions to translate statements to more primitive forms that can be worked with directly. Eventually the objects become familiar enough that they can be worked with directly and unpacking is no longer necessary, but until then we unpack.

For example a fraction $\frac{a}{b}$ is defined *indirectly* as a solution to an equation. Statements about fractions are unpacked by clearing denominators to remove the indirection, see steps (1) and (2) in the long–form solution above. This unpacking will be appropriate until after the exercises in §1.9.3, at which point it should be too routine to need explicit mention and can be "left to the reader".

1.6. **More about Inverses.**

[3]There is actually something still missing, see Problem 1.9.1.

1.6.1. *Significance.* Inverses may or may not make fractions uninteresting.

- There is not much point to decimal fractions like $\frac{4.209}{22.8888}$ because we can compute inverses (carry out the division).
- Exact real fractions like $\frac{1}{\pi}$ are useful because they retain a connection to the meaning of the number.
- The polynomial $(1-x)$ has inverse $\sum_{i=0}^{\infty} x^i$ in the formal power series ring. (Check to see that the product really is 1.) The fraction $\frac{1}{1-x}$ is usually more useful than the inverse. It defines an easily–computed function of x for $x \neq 1$, while the series form defines a function only for $|x| < 1$ and it is not easy to compute or work with.

1.6.2. *Problem: Inverses in Standard Rings.*

- Determine which numbers have inverses in the integers mod n.
- Show that a real–coefficient polynomial has an inverse in the ring of polynomials if and only if it is constant and nonzero.
- (Hard) Show that a formal power series has an inverse in the formal power series ring if and only if it has nonzero constant term. This extends the example given in the previous section.

1.7. **Difficulties with the Preliminary Version.** We return to the definition of fractions. The preliminary version given above has problems:

- We need to know there *is* a solution somewhere. It is usually not in the original context.
- We need to know there is *at most one* solution.

These are very different problems. It turns out that because the name maintains a connection to R we can almost ignore the existence problem. It does have to be addressed eventually, see §1.10.

If there is more than one solution then it it is hard to make sense of $\frac{a}{b}$ as a single thing. This has to be addressed immediately and we do that next.

1.8. **Zero divisors.** An element b in the ring is called a *zero divisor* if there is a $r \neq 0$ in R with $rb = 0$.

1.8.1. *Problem: Zero Divisors.*

- Show that if b is a zero divisor then there is an element a so $b\phi = a$ has more than one solution.
- Show conversely that if there is a with more than one solution then b is a zero divisor.
- Show that bc is not a zero divisor if and only if neither b nor c is a zero divisor. Hint: unpack using the conclusions just above.

The first two points imply that "non-zero-divisor" is exactly what we need for a fraction to make sense.

1.9. **Final version of the Definition.** If b is not a zero divisor then $\frac{a}{b}$ is a name for the solution of the equation $b\phi = a$. If b is a zero divisor then $\frac{a}{b}$ not defined.

Fractions with denominator a zero divisor, for example $\frac{3}{0}$, are undefined (mathematically illegal expressions) because they genuinely don't make sense. Trying to use them leads quickly to errors. For instance the definition requires that $\frac{3}{0} = \frac{4}{0}$, but this is a problem because $3 \neq 4$.

One can *force* division by zero to make sense, by use of an equivalence relation. Something like this is done in a slightly different context in the section on Grothendieck groups, §1.12. When applied to fractions the result is disappointing, see the third problem in §1.12.7. This gives another explanation of why we are stuck with the don't-divide-by-zero rule.

1.9.1. *Problem: Fractions and Zero Divisors.* We now see that the proof in §1.4 is incomplete: in order to be sure the fraction $\frac{a}{b}$ makes sense we must verify that if an element has an inverse then it is not a zero divisor. Rewrite both the long and short forms of the proof in §1.4 to include this.

1.9.2. *Problem: Zero Divisors in Standard Rings.*

(1) Show for examples (1), (3), (4), (5) in §1.2 that the only zero divisor is 0.
(2) Find an explicit form of the zero–divisor condition in the second example in terms of the modulus n. Compare this with the invertibility condition in §1.6.2.
(3) Laurent polynomials allow finitely many negative and positive exponents. Series allow infinitely many positive exponents. A natural generalization is series that are infinite in both positive and negative directions. However these rings are tricky to work with because they have a lot of zero divisors. As an example show that $(1-x)\sum_{i=-\infty}^{\infty} x^i = 0$. Generalize this: if r is a nonzero real number find a bi–infinite series whose product with $(r-x)$ is zero.

1.9.3. *Problem: Standard Fraction Facts.* Here the standard fraction facts are shown to hold for fractions in any commutative ring.

(1) Find a fraction expression for the sum $\frac{a}{b} + \frac{x}{y}$ (be sure to check the zero–divisor condition for the answer. Note that there is an implicit hypothesis that the fractions make sense: b and y are not zero divisors. Use this and Problem 1.8.1)
(2) Find a fraction expression for the product $\frac{a}{b}\frac{x}{y}$.
(3) Find a fraction expression for the fraction $\frac{a/b}{x/y}$. What is the condition required for this to make sense? (i.e. when is $\frac{x}{y}$ not a zero divisor?)
(4) Show that if c is not a zero divisor then $\frac{ca}{cb} = \frac{a}{b}$.

Since the rules are the same, people who can work accurately with integer fractions should also work with general fractions without explicitly referring to either rules or the definition.

1.10. **Rings of fractions.** We return to the existence problem mentioned in §1.7: if b is not a zero divisor in R then $\frac{1}{b}$ seems to make sense, but it is not an element of R unless b has an inverse. What, or where, is it? In fact we use fractions to define a new ring.

1.10.1. *Definition.* The *ring of fractions* of a commutative ring R, denoted here by $\text{Frax}(R)$, is the set of $\frac{a}{b}$ with b not a zero divisor, with two such being equivalent if they solve the same equation $b\phi = a$.

Equivalence means, for instance, that $\frac{a}{b}$ and $\frac{ac}{bc}$ are considered the same object even though they are different symbolic expressions.

The addition and multiplication formulas 1.9.3 are now no longer identities for preexisting objects but actually used to *define* addition and multiplication in the ring of fractions. Strictly speaking one should verify that these operations satisfy the rules required in a commutative ring. We will not do this because they follow easily and routinely from the rules in R.

1.10.2. *Examples.* We can now recognize some standard systems as being rings of fractions.

- The rational numbers are the ring of fractions of the integers.
- The rational functions[4] are the ring of fractions of the real polynomial ring.

1.10.3. *Problem: Fractions and Zero Divisors.*

- Show that $\frac{a}{b}$ is a zero divisor in the ring of fractions if and only if a is a zero divisor in the original ring.
- Show that every element in the ring of fractions that is not a zero divisor has an inverse.
- Show (conversely) that if every non-zero-divisor in R has an inverse, then the natural inclusion $R \subset \text{Frax}(R)$ is a bijection.
- Describe the ring of fractions of the integers mod n (see §1.9.2).

The natural inclusion in the third problem is defined by $a \mapsto \frac{a}{1}$. "Bijection" means that every element in $\text{Frax}(R)$ comes from exactly one element in R.

1.11. **Ring Homomorphisms.** To explore the relationship between rings and their rings of fractions we need a definition. A function $f: R \to S$ between two rings is said to be a *ring homomorphism* if it preserves the multiplication and addition operations and their units:

- $f(a+b) = f(a) + f(b)$
- $f(ab) = f(a)f(b)$, and
- $f(0) = 0,\ f(1) = 1$.

1.11.1. *Examples of Ring Homomorphisms.* Many of the standard relationships between the examples of rings in §1.2, for instance mod–n reduction going from the integers to the integers mod n, are ring homomorphisms. We add a few more:

- Fix a real number r. Show that evaluation at r defines a ring homomorphism from the polynomial ring to the real numbers.
- Show that the inclusion $R \subset \text{Frax}(R)$ of a commutative ring into its ring of fractions, defined by $a \mapsto \frac{a}{1}$, is a ring homomorphism.

1.11.2. *Naturality?* Applying a ring homomorphism f to an equation $b\phi = a$ gives $f(b)f(\phi) = f(a)$. Interpreting these as defining equations for fractions seems to show that f preserves fractions: $f(\frac{a}{b}) = \frac{f(a)}{f(b)}$. This should mean f induces a ring homomorphism on the rings of fractions: $\hat{f}: \text{Frax}(R) \to \text{Frax}(S)$. However:

- Find the error in this proposed construction.
- Find a really obvious ring homomorphism between two of the examples in §1.2 that does not extend to the rings of fractions.

[4] "Rational functions" really should be called "polynomial fractions". They are very useful as functions, but identifying them as elements in a ring of fractions is more fundamental. Note that the formal power series ring §1.2 Example (5) also has a ring of fractions but these generally cannot be interpreted as functions.

- Fix a real number r. Determine which polynomial fractions $\frac{p(x)}{q(x)}$ do *not* give a real fraction (and therefore not a real number) when evaluated at r.

1.12. **Grothendieck Groups.** Historically, fractions were first introduced as ratios, then used to encode and work with rational numbers. It was an unexpected bonus that they gave a way to mass–produce new rings by adjoining multiplicative inverses. The Grothendieck construction uses the same idea in a simpler context: *additive* systems without additive inverses.

The most familiar example is the natural numbers, and the construction adjoins additive inverses to produce the integers. There are many other examples but they require sophisticated preparation. This is why the additive version was described so much later than the fraction construction.

1.12.1. *Commutative Seimgroups.* A commutative semigroup is a set with a binary operation, denoted $+$, that is associative and commutative. Denote the set by N, then specifically:

- $a + b$ is defined for all a, b in N;
- $a + b = b + a$; and
- $a + (b + c) = (a + b) + c$.

There is a strong convention that an operation is entitled to be denoted "+" *only* if it has these properties. This means we can do arithmetic as usual with + operations, and don't have to explicitly think about the rules.

1.12.2. *Examples.*

- The natural numbers with the standard addition operation. This is denoted by $\mathbf{N}$.
- The natural numbers with operation given by minimum:

$$\min(a, b) = \begin{cases} a \text{ if } a \leq b \\ b \text{ if } b \leq a \end{cases}.$$

  This is sometimes called the *tropical* semigroup structure[5].
- If R is a commutative ring then R with the multiplication operation is a commutative semigroup. Caution: the use of multiplicative notation when thinking of it as a ring, and + for the same operation when thinking of it as a semigroup, is an endless source of confusion.
- The non-zero-divisors in a ring, again with multiplication as the operation, also form a semigroup. The third property in §1.8.1 is needed to see that this is true.

1.12.3. *Terminology.* A *group* is a set with a binary operation *with a unit element and inverses*. "Semigroup" weakens this by dropping the requirement that inverses exist. *Commutative* means that the operation is commutative, just as with rings. Mathematical experience suggests that the most fundamental object is a (possibly non–commutative) group. This therefore gets the short name, and related objects are described by modifying the name, just as with commutative rings (see §1.1.1).

[5]See the "tropical geometry" entry on Wikipedia.

1.12.4. *A Difficulty.* Recall that a fraction $\phi = \frac{a}{b}$ is defined to be the solution to the equation $b\phi = a$. We would like to similarly define a *difference* $\phi = (a - b)$ as the solution to the equation $b + \phi = a$, but there is a problem with this.

Recall that there are many solutions to the defining equation of a fraction if the denominator is a zero divisor. The solution was to only define fractions with non-zero-divisor denominators. The analog for addition is cancellation: $b + a = b + c$ implies that $a = c$. The strict analog of the fraction construction therefore only defines differences $a - b$ if b satisfies the cancellation condition.

We didn't mind having a condition on denominators of fractions. We do mind having a condition on negative objects (we want to be able to subtract without conditions). This requires a modification of the construction: requiring the defining equation to hold only after addition of the same term to each side[6]. Adding such a term is called "stabilization".

1.12.5. *The Construction.* Suppose $N, +$ is a commutative semigroup. $G(N, +)$ is defined to be equivalence classes of pairs of elements a, b, with equivalence classes written $[a - b]$:

- $[a - b]$ is equivalent to $[a' - b']$ if there is c so that $a + b' + c = a' + b + c$. (Think $a - b = a' - b'$, clear negative signs by adding b, b' to each side, then stabilize by c to avoid the cancellation problem).
- The operation $+$ is defined on equivalence classes by $[a - b] + [c - d] = [(a + c) - (b + d)]$.

1.12.6. *Problem: Identities in $G(N, +)$.*

- Show that $+$ is well–defined: if $[a - b] \simeq [a' - b']$ then $([a - b] + [c - d] \simeq [a' - b'] + [c' - d']$.
- Show that $[a - a] \simeq [b - b]$ for all a, b, and that this equivalence class is a unit for the operation: $[a - b] + [c - c] \simeq [a - b]$. We follow tradition by denoting the equivalence class $[a - a]$ by 0.
- Show that $[b - a]$ is an additive inverse for $[b - a]$.

The outcome is that $G(N, +)$ is a commutative group (i.e. commutative semigroup *with inverses*). The construction adjoins inverses.

1.12.7. *Problem: Examples.*

- Show that the function $[a - b] \mapsto a - b$ defines an isomorphism[7] from the Grothendieck group of the natural numbers with standard addition, to the integers.
- Describe the Grothendieck group of natural numbers with the tropical (min) operation.
- Suppose R is a commutative ring. Show that the Grothendieck group of R, with multiplication as the operation, is trivial (everything equal to the identity element). This semigroup is an example in §1.12.2 where there is a warning about notation problems with the operation. The best way to proceed is to translate the definition of $G(R, \times)$ into multiplicative notation, and see it as a modification of the definition of fractions.

[6]The analogous modification of the fraction construction is to allow multiplication by an arbitrary element. The effect is to enable division by anything, including zero divisors if there are any. The next-to-last example in §1.12.7 illustrates what happens.

[7]"Isomorphism" here means one-to-one and onto, and takes $+$ in one group to $+$ in the other.

- Describe the Grothendieck group of the *nonzero* integers, with multiplication as operation.

Generalize the example just above to show:

1.12.8. *Proposition.* Suppose R is a commutative ring. Then the group of invertible elements in the ring of fractions of R is the Grothendieck group of the semigroup of non-zero-divisors of R.

After unwinding the definitions you should see this as essentially obvious. The impressive–sounding statement is the result of having two terminologies for essentially the same construction.

2. Area

The object is to explore a formula for areas of polygonal figures in the plane, using coordinates of the vertices.

2.1. **Polygonal closed paths.** Suppose $(p_0, p_1, \ldots, p_n)$ are points in the plane $\mathbf{R}^2$. The *oriented closed path* with these as vertices is obtained by joining p_i to p_{i+1} for $0 \le i \le n$ where, if $i = n$, we set $p_{n+1} = p_0$.

Closed means it goes back to the starting point: this is the effect of the $p_{n+1} = p_0$ convention. *Oriented* refers to the preferred direction on the path coming from the order of the vertices. Problem: Draw[8] a few of these by choosing points at random, numbering them, and then connecting them, drawing arrows to indicate direction on the edges. Do one with only two vertices.

2.2. **The Project.** Suppose P is an oriented closed path with vertices $p_0, \ldots p_n$. Denote the coordinates by $p_i = (x_i, y_i)$ and define

$$A(P) = \frac{1}{2}\sum_{i=0}^{n}(x_i y_{i+1} - x_{i+1} y_i). \tag{1}$$

The project is to show that if P is a closed polygonal path then $A(P)$ is the area enclosed by P. More precisely, find conditions under which this is true.

2.2.1. *Problem: Example.* If this works then it makes areas easy to compute when coordinates of the vertices are known. Find $A(P(t))$ for quadrilaterals $P(t) = ((0,0),(1,t^2),(1-t,1-t),(t^2,1))$ when $-1 \le t \le 1$. Draw a few of these to see what is happening in this family. Determine the t at which $A(P(t))$ attains it's maximum.

2.2.2. *Notes.*

- This may seem unlikely: how can something that only uses the path give the area? Also, the definition of "path" allows self–intersections so does "enclosed" even make sense? Or is it just for simple closed paths?
- The statement can't be right even for simple closed curves: areas should be nonnegative, but $A(P)$ can be negative: reversing the order of the vertices in P reverses the sign of $A(P)$. This has to be sorted out.

[8] Draw on paper, with a pen or pencil. There is something about actual drawing that significantly aids learning, and students really are expected to do this.

- It is useful to observe that A is defined for curves that aren't closed. In fact it is defined for a single edge, and

$$A(P) = \sum_{e_i \text{ edges of } P} A(e_i) \tag{2}$$

2.2.3. *Problem: Test Cases.* To clarify what is going on, try some special cases (with pictures!). Compute both A and the area.

- A triangle[9] with p_0 at the origin and p_1 on the positive x axis. This has vertices $((0,0),(x_1,0),(x_2,y_2))$ with $x_1 > 0$. Note there are several cases depending on whether (x_2,y_2) is above, on, or below the x axis.
- A trapezoid (4 vertices) with p_0, p_1 as above and the segment from p_2 to p_3 horizontal.

You should see there is a sign problem: the definition needed to get the sign right is *counterclockwise orientation of the boundary.* Essentially it means that if you imagine yourself moving along the boundary in the direction specified by the orientation then the "inside" of the region is on your left[10].

2.3. **A Difficulty, and a Strategy.** Area is not defined in school mathematics. Students are taught formulas for areas of simple figures, but these are obtained from basic examples (especially rectangles) and justified heuristically. Without a definition, or at least a reasonably general way to compute, there is almost nothing to connect with. How can we expect to relate A to area under these circumstances?

In fact we see that this is a problem in the usual development of mathematics. The first real definition of area is given in multivariable calculus: the area of a region is the double integral of the function 1 over the region. This definition makes many calculations easy, and gives the formula $A(P)$ via Stoke's theorem. This is connected with earlier work by showing it gives the familiar answers for triangles, circles, etc. It is *not* shown to agree with an earlier *definition* of area because, of course, there wasn't one. In essence, earlier work becomes obsolete and is discarded.

Our objective is still to connect A to area, and to do this without going through calculus. The plan is to list properties that area—however it is defined—should have. We then verify that the function A has these properties. If the properties are strong enough to completely determine the area of a polygonal region it will follow that A must be area.

2.3.1. *Properties of Area.* Area of polygonal regions should satisfy:

Invariance under rigid motions: Rotation and translation do not change area;

Additivity: if a region is split into two pieces then the total area is the sum of the areas of the pieces; and

Standard triangles: areas of triangles with one edge on the x axis is one-half (length of base) times (height).

[9] Do this symbolically. Do *not* put numbers in for x_1, etc.

[10] This is imprecise but good enough for the present. See §2.6 for a precise version.

A list like this is always a bit dangerous: we are working blind, and we might assume more than is actually true. In that case deductions made from the assumed properties will lead to a contradiction and the whole effort will collapse. To minimize risk we try to get the job done with the weakest possible assumptions.

Eventually, by using A, we will see that area has many additional properties that we would not dare to assume when working blind.

2.3.2. *Rough Argument for Sufficiency.* We suggest why these properties should be enough to determine the area of a region. It should be possible to divide a polygonal region into triangles. Using additivity the whole area is the sum of the areas of these triangles, so we only have to show that areas of triangles are determined. Any triangle can be translated so one vertex is at the origin, and then rotated so another vertex (and therefore an edge) is on the x axis. Since area is unchanged by rigid motions, the area of general triangles is determined by areas of these special cases. But areas in these cases is specified in the standard–triangle property.

This argument is not solid. The part about determining area of general triangles is OK (i.e. essentially a proof), but subdivision of regions into triangles needs to be done carefully. There might be a subtle difficulty with this that could require restriction to special polygonal regions, for instance convex ones. However the proper next step is to see if A has the properties. The reason is that if this fails then we can conclude that A does not give areas, or the whole approach has to be modified, and we don't have to worry about the subdivision argument. If A passes then we can return to the subdivision argument.

2.4. **Properties of A.** Terminology used in the problems is explained in the following section. After that we compare with to determine exactly what remains to be done to complete the project.

2.4.1. *Problem: Properties of A.* Suppose P is a polygonal region with vertices p_i for $i = 0, \ldots, n$, and $p_{n+1} = p_0$. Prove the following:

- **Translation Invariance:** If q is any point in the plane then the region $P+q$ with vertices $p_0 + q, p_1 + q, \ldots$ satisfies $A(P) = A(P + q)$.
- **Matrix Transformations:** If R is a 2×2 matrix then $A(RP) = \det(R)\, A(P)$. See Notes below.
- **Additivity:** Suppose Q is a polygonal path beginning and ending at vertices of P, see Notes below. Denote the two regions obtained by splitting P along Q by PQ_1 and PQ_2. Then $A(PQ_1) + A(PQ_2) = A(P)$.
- **Subdivision:** Suppose q is a point on the edge between p_i and $p[+i+1$. Denote by P_q the region defined by inserting q in the vertex list between p_i and $p[+i+1$. Then $A(P) = A(P_q)$.
- **Cyclic Permutation:** The region obtained by cyclically permuting the vertices, i.e. $(p_i, \ldots, p_n, p_0, p_1, \ldots, p_{i-1})$, has the same A value.

2.4.2. *Notes.*

- In Rotation Invariance, a 2×2 matrix operates on points in the plane by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)(x, y) = (ax + by, cx + dy)$. In particular, rotations do not change A because they have determinant 1. There are matrices like the shear transformations $\left(\begin{smallmatrix} 1 & r \\ 0 & 1 \end{smallmatrix}\right)$ that are not rotations but have determinant 1 and so also don't change A. When we know that A is area it will follow that area is not changed by these either.

- In Additivity, Q is of the form $(p_i, q_1, \ldots, q_k, p_j)$. Suppose $i < j$. Then the split regions are $(p_i, \ldots, p_j, q_k, \ldots, q_1)$ and $(p_j, \ldots, p_n, p_0, \ldots, p_i, q_1, \ldots, q_k)$. What happens if $j < i$ or $j = i$? Hint: use the edge–sum description in equation (2).
- There are two points to the Subdivision and Cyclic Permutation properties. First, A is defined using a specific ordering of specific vertices. These properties show that A depends on the underlying geometric figure (and the direction on the boundary), not specific vertices. These properties will also be needed to divide regions into triangles.

2.4.3. *Taking Stock.* The function A has all the properties we could want, and more: additivity works even if the cutting curve intersects original, or if it lies *outside* the original region. This strange behavior has something to do with it being able to take negative values. In any case the precise relation to area is still unclear.

Referring back to the discussion in §2.3.2, we see that all the pieces are in place except for the argument about cutting a region up into triangles. In particular A is now known to give area of positively oriented triangles. Evidently the cutting argument is where the negative–value and crossing problems get sorted out. This means a logically complete version of the argument will have to be fairly complicated.

Draw some pictures to explore what can go wrong. The curve could be a wild scribble, or have lots of sharp points, or wind back and fourth like a maze puzzle, or all of these. It is hard to imagine a cutting strategy that would do a good thing in all these cases.

Instead of trying to find a strategy for cutting a region into smaller pieces we make it bigger. It is relatively easy to prove the result for convex polygons. The general simple closed case can be done by filling in concave areas until it becomes convex.

2.5. **Orientations and Convex Polygons.** We develop ideas that lead to definitions of both convexity and orientation. These are used to prove that A gives the area of the region enclosed by a positively oriented *simple* (i.e. no self–intersections) polygon. The actual result is weaker than it sounds because we don't have a good definition for "region enclosed". Eventually this will be provided by winding numbers but here we use a more elementary, and much more cumbersome, approach.

The standard definition of a convex *region* is that if two points are in the region then the straight line between them is also in the region. Here we would want to apply this to the region enclosed by the curve. However we can't do that because—again—we don't have a criterion for which points are inside and which are outside. Again we use an approach special to the situation.

2.5.1. *Extremal Vertices and Edges.* Suppose P is a closed polygonal path. A vertex is called *extremal* if there is a bi–infinite straight line that intersects the polygon in exactly this vertex. A line is called an *extremal line* if it intersects the polygon in at least two points and there are no vertices on one side of the line.

For example suppose y_i is the smallest y coordinate of any vertex. The horizontal line (t, y_i) for $t \in \mathbf{R}$ passes through the vertex p_i and there are no vertices below this line. If p_i is the only vertex on the line then p_i is an extremal vertex. If there are other vertices on the line then this is an extremal line.

The outermost vertices on an extremal line are extremal vertices. For instance take the lowest vertex. Rotate the line about this vertex toward the side with no vertices. This moves the line away from the later vertices, but does not immediately introduce new intersections. The slightly rotated line therefore identifies the vertex as extremal.

2.5.2. *Problem: Example.* Draw a complicated, jagged, closed polygon. Locate the extremal vertices and draw the extremal lines with a ruler.

2.5.3. *The Convex Hull.* A new polygon can be constructed from the extremal vertices by ordering them so adjacent ones lie on an extremal line. This is called the *convex hull* of the polygon, and is the smallest convex set containing the polygon. There are two orderings that do this (up to cyclic permutation) and we want to specify one as being "positive".

Obtain particular extremal lines by translating and rotating the x axis so that the polygon lies in the translated upper half plane (nonnegative y coordinates). Each such extremal line contains two extremal vertices: the ones corresponding to minimal and maximal x coordinate. Order the vertices by placing the one with maximal x coordinate just after the minimal one. Then this is the positive (or counterclockwise) orientation of the convex hull.

Referring back to the test cases in §2.2.3, we see that all the vertices on the triangles and quadrilaterals there are extremal. The positively oriented examples are the ones with nonnegative y coordinates, while the others are negatively oriented (opposite of positive). Note that this definition of orientation exactly determines the sign of A for these examples. In particular if P is a positively oriented triangle then $A(P)$ is the area of P.

2.5.4. *Problem: Areas of Convex Polygons.* Define a polygon to be *convex* if all vertices are extremal. Show that if P is a positively oriented convex closed polygon then $A(P)$ is the area of the region enclosed by P.

To do this choose a vertex (say p_0) and consider the triangles (p_0, p_i, p_{i+1}). Show that each of these is positively oriented and their union is P. Then use additivity.

2.6. **Simple Closed Polygons.** A closed polygon is called *simple* if no edge intersects another edge or vertex except at it's endpoints.

We can define "positive orientation" for simple polygons: suppose e_i and e_{i+1} are adjacent extremal vertices, ordered using the positive orientation of the convex hull. Denote the line through these points by E. P splits into two polygonal paths from e_i to e_{i+1}, both contained in the half–plane on one side of E. Since there are no self–intersections in P one of these is contained inside the region bounded by the other and E. We define P to be positively oriented if e_i is before e_{i+1} in the ordering on the *inner* path.

We postpone the argument that there really is an "inner" and "outer" path. It is also true that the inner path cannot contain any other extremal vertices (a line through a vertex has to intersect the outer curve). If P is positively oriented then this implies the ordering on the vertices of P gives the natural positive order on the extremal vertices, and this in turn implies that the definition of positive does not depend on which pair of extremal vertices are used. All these things can be proved with the techniques at hand, but somewhat awkwardly. A short clear proof will be possible after winding numbers are developed.

2.6.1. *Proposition: Areas of Simple Polygons.* Show that if P is a simple closed polygon then the area of P is $(-1)^k A(P)$, where $k = 0$ if P is positively oriented, $k = 1$ otherwise.

It is sufficient to show that $A(P)$ is the area if P is positively oriented, because reversing orientation changes the sign of A.

Proceed by induction on the number of vertices of P, starting with triangles ($n = 3$). The induction step is: suppose the proposition is true for all simple polygons with fewer than n vertices, and suppose P has exactly n. Then show that the proposition is also true for P. Note that if P is not convex then there are adjacent extremal vertices e_i and e_{i+1} so that the edge between them is not a union of edges of P. There is a path in P between these vertices that does not contain any other extremal edges. Define Q by replacing this path in P with the direct edge between e_i and e_{i+1}. Observe that Q has fewer vertices than P, and argue that P is obtained from Q by cutting along the replaced path.

2.7. **Winding Numbers.** Winding numbers count how many times a curve in the plane "goes around" various points. Winding numbers are the starting point for remarkable developments in analysis, topology and other areas, and provide a complete understanding of how A relates to area.

The rough idea is that a general polygonal curve (with intersections) does, in a sense, enclose a region. But this region can overlap itself, and—due to folding—can have pieces with negative orientation and therefore negative "area". The function A gives the area when counted with multiplicity of overlaps, and winding numbers give the multiplicity. The formal statement is given in §2.7.5, but properties of winding numbers must be developed first.

2.7.1. *Definition.* Suppose P is a polygonal path, a is a point not on P, and R is a ray beginning at a that does not pass through any of the vertices of P. The *winding number* of P around a is the number of intersection points of R with edges of P, *counted with signs*.

Signs are assigned to intersection points as follows. Represent the ray as obtained from the positive x axis by translation and rotation. The translation and rotation act on the whole plane, and we can define the *positive side* of the ray to be the image of the upper half plane. Next recall that edges have a preferred direction coming from the order on the vertices. Define an intersection point to have sign $+1$ if the edge direction goes from the negative to the positive side of the ray. Sign -1 corresponds to a positive-to-negative crossing.

Note there is an intersection associated with each crossing edge. It might be that multiple edges cross at the same point on R, in which case this point is counted multiple times.

2.7.2. *Problem: Draw an Example.* Let R be the positive x axis, thought of as a ray beginning at 0. Draw a complicated closed curve that winds around 0 but with occasional changes in direction so some crossings of R are positive and others are negative. Then trace along the curve in the preferred direction and record the direction at each crossing point with an arrow. Add to determine the winding number.

2.7.3. *Problem: Winding Numbers are Well–Defined.* Show:

(1) Any ray starting at a and missing the vertices of P gives the same winding number.
(2) If a and b are joined by a polygonal path disjoint from P then P has the same winding numbers around a and b.

These are consequences of:

Lemma: suppose $R(t)$ is a 1–parameter family of rays with origins disjoint from P. Then $R(0)$ and $R(1)$ give the same winding number.

In part (1) of the problem, a 1–parameter family can be obtained by rotating about the point a. In part (2), start with a ray at a and use translations to move the whole ray so that the endpoint moves along the path. Caution: the statements in the problem are a bit imprecise. Be very precise about what these arguments actually prove, and show how they fit together to give the conclusion.

To prove the lemma, *assume* that there are only finitely many values of t for which $R(t)$ contains vertices of P. For a random 1–parameter family this may not be true. It is true for the families actually used (rotations and translations) and can be arranged without much difficulty in general, but there is not much benefit to going into detail here. Then:

- Argue that on an interval of t values that does not contain a vertex intersection, the number doesn't change. This should be easy.
- Argue that the number also doesn't change when the parameter passes through a value with vertex intersections. Do this with pictures of the various ways edges can enter and exit a vertex, and how a ray could sweep through the picture. Don't get too formal, but be sure you have all possibilities represented. Caution: the ray could contain an edge of P!

2.7.4. *The Polygonal Jordan Curve Theorem.* The polygonal path property in part (2) of the problem above inspires the following definition: A region is said to be *connected* if any two points can be joined by a polygonal path. When applied to complementary regions this means a path disjoint from the original curve.

Problem. Show that a *simple* closed polygonal curve divides the plane into exactly two connected regions:

- a (unbounded) region in which the winding number is 0; and
- a region in which the winding number is 1 if P is positively oriented, -1 otherwise.

To prove this construct polygonal paths by going along a ray to the first intersection point with P, then following along beside P to an intersection point with another ray. Evaluate the winding number by starting near an extremal vertex.

The general Jordan Curve Theorem asserts that a *continuous* simple closed curve divides the plane into two connected regions. The proof for continuous curves is much more difficult than the polygonal case.

2.7.5. *Theorem.* A closed polygonal curve P divides the plane into a finite number of connected subregions. $A(P)$ is the sum over these, of the area of the subregion times the winding number of P around a point in the subregion.

According to part (2) of Problem 2.7.3, the winding number is constant on a connected region, so it doesn't matter which point in the subregion is used.

To organize the proof, introduce the notation $W(P)$ for the weighted area sum.

- Both $A(P)$ and $W(P)$ are defined for *collections* of closed curves, not just single closed curves. We use this in a cutting argument that splits curves into pieces. (Another approach is given in §3.3).
- Both are additive with respect to cutting and unions of multiple curves.
- By quoting previous results we can conclude that they are the same for simple closed P.

There are several ways to complete the proof from this point. One approach is by induction on the number of complementary regions, for a single closed curve. The induction starts with two complementary regions (the simple closed case) because we know A and W agree for these.

For the induction step suppose the statement is known for n or fewer regions, and suppose P has $n+1$. Consider a segment of P between two self–intersection points. Cut along this segment to convert P into two closed curves. Together these have the same complementary regions as P (because the segment is in P), and winding numbers for the union are the same as for P. Therefore neither A nor W is changed by this. However we now have two pieces which must both be "smaller" than P, so the induction hypothesis applies to both pieces.

This needs refinement. If P traces over itself several times then the cutting procedure can give pieces with the same complementary regions. To fix this we need something more subtle than a simple count of regions. Try using a weighted count: the sum of the absolute values of the winding numbers. If that doesn't work try inducting on the number of self–intersections. However beware that it may trace over itself somewhere, so have segments of self–intersection and not just isolated points.

3. More About Area

This section gives explorations and elaborations.

3.1. **Area and Rings.**

In this section and the next we try to get more elaborate versions of area by using the same formula in different rings. One works, one doesn't.

3.1.1. *Complex Area?*

Think of the plane $\mathbf{R}^2$ as complex numbers, with $p = (x, y)$ thought of as $p = x + iy$.

Now suppose $p_k = x_k + iy_k$ are vertices in a polyhedral path. The cross term $x_k y_{k+1} - x_{k+1} y_k$ in the definition of $A(P)$ is very nearly the imaginary part of the complex product $p_k p_{k+1}$. To get the sign on the second term right, recall that complex conjugation is defined by $\overline{x + iy} = x - iy$. Then $x_k y_{k+1} - x_{k+1} y_k = \mathrm{Im}(\overline{p_k} p_{k+1})$. Define

$$A_{cx}(P) = \sum_{k=0}^{n} \overline{p_k} p_{k+1},$$

then the previous definition is the imaginary part, $A(P) = \mathrm{Im}(A_{cx}(P))$.

Problem Show that the real part of $A_{cx}(P)$ is not invariant under some area-preserving transformations (i.e find one). Consequently A_{cx} does not qualify as a generalized area.

3.1.2. *Polygons in Motion?* Suppose we have a one–parameter family of polygonal curves $P(t) = (p_0(t), \dots, p_n(t))$ defined for t in an interval (a, b). The objective is to apply the formula defining A to both coordinates and derivatives of coordinates, to see if this encodes something useful. The first step is to construct a ring where the invariant will be defined.

The ring $\mathbf{R}[\delta]/(\delta^2 = 0)$ Extend the real numbers by adjoining δ with $\delta^2 = 0$ rather than $i^2 = -1$ as in the complex numbers[11]. More precisely, the multiplication in this ring is given by

$$(a + \delta b)(x + \delta y) = ax + \delta(ay + bx).$$

Denote this ring by $\mathbf{R}[\delta]/(\delta^2 = 0)$.

Problem: Zero Divisors Find the zero divisors (see §1.8) in $\mathbf{R}[\delta]/(\delta^2 = 0)$. We know there is one because we put it there ($delta^2 = 0$).

Encoding Derivatives Suppose $f(t)$ is a real–valued function defined and differentiable on an interval (a, b). Define $f^\delta \colon (a, b) \to \mathbf{R}[\delta]/(\delta^2 = 0)$ by $f^\delta(t) = f(t) + \delta Df(t)$, where Df denotes the derivative.

Problem: Products. Show that if f, g are both defined and differentiable on (a, b) then $(f \times g)^\delta = f^\delta \times g^\delta$.

Clarification of notation: $(f \times g)^\delta$ is the δ construction applied to the ordinary product of real–valued functions. $f^\delta \times g^\delta$ is the product in the ring $\mathbf{R}[\delta]/(\delta^2 = 0)$. Multiplication is written explicitly (i.e $f \times g$ rather than fg) to avoid confusion with composition of functions. Unpack carefully.

Definition of A^δ. Now suppose $P(t) = (p_0(t), \dots, p_n(t))$ is a family of polygons defined for t in (a, b), and suppose that all the coordinate functions are differentiable. Define $A^\delta(P)$ by using the formula (1) used to define A, but in the ring $\mathbf{R}[\delta]/(\delta^2 = 0)$ using the extended coordinate functions $x_i^\delta(t)$, $y_i^\delta(t)$.

Problem: Example. Find $A^\delta(P)$ (as a function of t) for the example in §2.2.1. As an intermediate step write out the extended polygon $P^\delta(t)$.

Problem: Describe A^δ. Show that if $P(t)$ is a differentiable one–parameter family of polygons then

$$A^\delta(P)(t) = A(P(t)) + \delta D(A(P))(t).$$

In words, this means that when we work in the ring $\mathbf{R}[\delta]/(\delta^2 = 0)$ the formula for A gives the derivative of the area as well as the area itself.

3.2. **Area and Vectors.** Here we shift point of view and express a polygonal path as a pair of vectors. See §**??** for an account of where this came from.

Instead of thinking of a sequence of pairs of numbers $(x_1, y_1), \dots$ we could think of it first as a $n \times 2$ matrix

$$\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{pmatrix}$$

[11] The complex number i is sometimes written $\sqrt{-1}$. Could the "number" δ be written $\sqrt{0}$?

and then as two vectors $X = (x_1, x_2, \dots)$ and $Y = (y_1, y_2, \dots)$ of the same length. These should be thought of as column vectors ($n \times 1$ matrices) but to save space they are usually written as rows.

3.2.1. *A as a Product.* Separate the two terms in the definition of A, 1, shift the index of one, recombine and factor out the x term:

$$\begin{aligned} 1/2\sum_{i=0}^{n}(x_i y_{i+1} - x_{i+1} y_i) &= 1/2(\textstyle\sum_{i=0}^{n} x_i y_{i+1} - \sum_{i=0}^{n} x_{i+1} y_i) \\ &= 1/2(\textstyle\sum_{i=0}^{n} x_i y_{i+1} - \sum_{i=0}^{n} x_i y_{i-1}) \\ &= 1/2\textstyle\sum_{i=0}^{n}(x_i(y_{i+1} - y_{i-1}) \end{aligned}$$

The last expression is a dot product of the vector X and something obtained by shifting the coordinates of Y. To express this last we need a notation.

Define the *Right Rotation* of a vector by $R(y_1, y_2, \dots, y_n) = (y_n, y_1, y_2, \dots, y_{n-1})$. This is represented by multiplication by the $n \times n$ matrix with 1s just below the diagonal, a 1 in the upper right corner, and all other entries 0. We therefore usually write the function as a matrix product (indicated by a dot), $R \cdot Y$. Again note Y is considered as an $n \times 1$ matrix.

Problem Show that the left rotation R^{-1} is the transpose R^t.

We can therefore write the vector with i^{th} coordinate $y_{i+1} - y_{i-1}$ as $(R^{-1} - R) \cdot Y$. With all this in hand we get

$$A(X, Y) = X^t \cdot (R^{-1} - R) \cdot Y$$

where $A(X, Y)$ denotes (no surprise) the function A applied to the polygonal path with the indicated vertices.

3.2.2. *Caution about the Transposition of X.* The X is transposed to make it a $1 \times n$ matrix so the product is defined. This is quite important. For example if we rotate both X and Y one place to the right then we get the same path but starting at a different vertex. The area is unchanged and this should be visible from the formula. Begin with routine expansion of the definitions:

$$\begin{aligned} A(R \cdot X, R \cdot Y) &= (R \cdot X)^t \cdot (R^{-1} - R) \cdot R \cdot Y \\ &= X^t \cdot R^t \cdot (R^{-1} - R) \cdot R \cdot Y \end{aligned}$$

But $R^t = R^{-1}$ so the middle terms reduce to $(R^{-1} - R)$ and we get the expression for $A(X, Y)$.

There is a caution about this argument. We have both R and R^{-1} in the expression and the idea is to cancel them. *However* matrix multiplication usually doesn't commute so we can't move the R^{-1} past the $(R^{-1} - R)$ just by general principles. You can see that the formula works anyway by multiplying it all out.

There is a shortcut: it is easily seen that powers of a single matrix (including negative powers if they exist) *do* commute with each other. In this case only powers of R are involved so in fact the product $R^{-1} \cdot (R^{-1} - R)$ *does* commute. Therefore in this case we can just commute the R^{-1} and R to be adjacent and cancel them. It is simpler to see the identity this way, and this sort of thing will be useful when we have to deal with bigger products.

Warning: if you use this fact you *must* say "because powers of a matrix commute...". Otherwise it will look like a common error (forgetting matrix multiplication almost never commutes) and *should* be counted wrong: this error is so dangerous it must be caught every time.

The point demonstrated here is that we get cancellation because the transpose on X^t changes R to R^{-1}. It also reverses the order of the product (this is a property of transposition and is *not* commuting). If we had been sloppy and omitted the transpose then the expression would be wrong. The biggest danger with something like this is that it is often not clear what the problem is. If we can't find anything interesting, is it because there is nothing interesting to find, or because an incorrect expression can't find it? This is less trouble for students because they can check the answer, but it is a constant hazard for mathematicians. See §**??**.

3.2.3. *Area–Killing Vectors.* The question here is: are there "area–killing vectors" Y so that $A(X, Y) = 0$ for *all* X?

The standard inner product $X^t \cdot Y$ is "nondegenerate" in the sense that $X^t \cdot Z = 0$ for *all* X only if $Z = 0$. Therefore the killer vectors are ones with $(R^{-1} - R) \cdot Y = 0$, or in other words with $R^{-1} \cdot Y = R \cdot Y$.

- Interpret this as a relationship between entries of Y.
- Show that if Y has *odd* length then all entries of Y are the same. If all y coordinates of points in P are the same then they lie on a horizontal line. It is certainly reasonable that the area should be 0 no matter what the x coordinates.
- Argue, using invariance under rotation, that any P whose vertices lie on a line must have zero area.
- Figure out what $R^{-1} \cdot Y = R \cdot Y$ implies if Y has *even* length.
- Interpret the result in terms of horizontal lines, then get a more general statement by rotation.
- I didn't expect this outcome, and had to draw some pictures to feel comfortable with it. Do this, starting with Y of length 4.

3.2.4. *Bilinearity, and Midpoint Polygons.* An immediate consequence of the product formulation is that $A(X, Y)$ is linear in both X and Y. Explicitly this means

$$A(r_0 X_0 + r_1 X_1, Y) = r_0 A(X_0, Y) + r_1 A(X_1, Y)$$

(real coefficients r_i) and similarly for Y. If we take combinations of X_0, X_1 and of Y_0, Y_1 we get a 4–parameter family of polygons, and a formula for their areas in terms of the four basic ones ((X_0, Y_0), (X_0, Y_1) etc.). This is a little too free–form to be really interesting, so we explore a classical formula where the various polygons are closely related.

Suppose $P = (X, Y)$ is a polygonal path. The *midpoint polygon* of P has vertices the midpoints of the edges. Explicitly, the i^{th} vertex is $\frac{1}{2}(\,(x_i + x_{i-1}, y_i + y_{i-1})\,)$. In vector notation this becomes $\frac{1}{2}(X + R \cdot X, Y + R \cdot Y)$. We can also write $X + R \cdot X$ as $(I + R) \cdot X$, where I is the identity matrix.

(1) Show that A of the midpoint polygon is given (up to a constant) by

$$X^t \cdot (I + R^{-1}) \cdot (R^{-1} - R) \cdot (I + R) \cdot Y.$$

In particular find the constant, and justify the form of the first R term.

(2) Show that the product of the R terms in this expression is equal to

$$2(R^{-1} - R) + (R^{-2} - R^2).$$

Rather than multiply it all out, commute the first and second terms (*don't* forget to write the magic words for this) and just multiply out the first and third.

(3) Interpret $X^t \cdot (R^{-2} - R^2) \cdot Y$ as A of a polygon with the same vertices as the original, but used in a different order. (This is called the "skip" polygon. If there are an even number of vertices it is actually two polygons.).
(4) Put these together to get a formula for the area of the midpoint polygon in terms of the areas of the original and the skip polygon.

Is the use of the midpoint (i.e. coefficients $\frac{1}{2}$) really essential in this formula? Try other combinations of the polygon and its rotation:

(1) Expand A of $r(X, Y) + s(R \cdot X, R \cdot Y)$, and simplify as above.
(2) For which values of r, s can this be expressed in terms of the areas of the original polygon and the skip polygon, and what is the expression?

3.2.5. *Zero–Area Skip Polygons.* In §3.2.3 we found criteria for polygons to have zero area. We use the coordinate–free versions, e.g. "all vertices lie on a line".

(1) Use these to find conditions on a polygon that ensure the associated skip polygon has zero area.
(2) The most interesting case is when the number of vertices is divisible by 4, though the case $n = 4$ is pretty trivial. Use the criterion to draw some polygons with 8 vertices with zero–area skip polygons.

3.2.6. *Morphing.* We want to "morph" P to $R(P)$ without changing the area. Then the formula at the end of §3.2.4 for linear combinations is

$$A(rP + sR(P)) = (r^2 + s^2)A(P) + rs\,A(\text{skip}P) \tag{3}$$

If we want this equal to $A(P)$ it becomes

$$(r^2 + s^2)A + rs\,A_{\text{skip}} = A$$

and the question is whether we can go from $(r, s) = (1, 0)$ to $(0, 1)$ in the set of (r, s) that satisfy this equation.

This breaks into two cases: $A = 0$ and $A \neq 0$.

If $A = 0$ we need $r = 0$ or $s = 0$, so the solution set is the union of the coordinate axes. We can go from $(1, 0)$ to $(0, 1)$ in this set, but note we have to go through $(0, 0)$ to do it. In other words the morph shrinks P to a point and then expands it back out to $R(P)$.

Now suppose $A \neq 0$, and define $k = A_{\text{skip}}/A$. The equation becomes:

$$(r^2 + s^2) + k\,rs = 1$$

Problem. Describe this solution set for various values of k, and in particular find values where qualitative behavior changes:

(1) For which k are there two different paths from $(1, 0)$ to $(0, 1)$?
(2) Is there a value for which there is no morphing path?
(3) Are there any restrictions on the values of k coming from actual polygons?

The last question requires constructing examples. It is not worth spending a whole lot of time on, so skip it if you don't see what to do reasonably quickly.

If you have access to software that displays polygons, do the following:

(1) Explicitly parameterize part of the solution set to get a morphing path (Use polar coordinates and parameterize by the angle from the x axis).
(2) Watch various examples morph and see if you can qualitatively describe any of the behavior, e.g. in terms of k.

(3) You will see that some examples move around, and that the movement depends on location (translate to see this). The reason is that the polygons that can be obtained by linear combination depend on position, so these morphs go through different spaces. It would take us too far afield to explore this but two directions might be:
   - Is there a special location with a nice morph (something like "centroid at the origin")? This probably wouldn't work if it is possible for polygons and their skip polygons to have different "centroids".
   - What happens if we enlarge the space available for morphing to something like $rP + sR(P) + (u, v)$ with four free parameters?

3.3. **Folding.** [[unfinished]]

4. Comments for Students

4.1. **Route to a Formula.** The Area and Vectors material in §3.2 is not deep but I stumbled on it accidently and inefficiently. The story may be reassuring to students.

I heard about the midpoint polygon area formula in a lecture by Thomas Banchoff on use of dynamic geometry software. He illustrated how students might test and (laboriously) prove the result geometrically for convex polygons. It seemed a good exercise for the approach here, though I have since learned that in his course he also gave an analytic treatment similar to this one.

First I wrote out the expression (1) explicitly using the average description of the midpoints. The algebra was a mess but it did work out. I also tried it for other points on the edges: things of the form $tx_i + (1-t)x_{i-1}$, not just the midpoint $t = \frac{1}{2}$. This worked too, though it took a while to get the algebra right.

I started looking for a better way to organize it. It requires combining x and y coordinates in different ways, which seemed unnatural geometrically. One could (using rotations) think of the coordinates as two orthogonal projections rather than the standard coordinates. The complex numbers have specific orthogonal projections built in, via complex conjugation. I thought this might be a clue that the complex formulation in §3.1.1 might be a good setting. I spent a fair amount of time trying to see something good in this, and failed. I might have quit too soon, though, see [2].

I went back to looking for patterns in the sum expression. I had previously noticed the reorganization and factoring used in §3.2.1 but hadn't thought anything of it because I assumed it would be a bad idea to separate the coordinates (e.g. into vectors X and Y). This time I knew separation had to be part of the story. This freed me to recognize the expression as a dot product. It did not take long to get the tidy description using matrix products and the rotation matrices.

The matrix formulation showed that A defines an anti–symmetric bilinear function (anti–symmetric means $A(X,Y) = -A(Y,X)$, which happens here essentially because the orientation gets reversed). A standard question about such things is degeneracy: how many "area–killing" Y there are, in the sense of §3.2.3. Describing solutions of $(R^{-1} - R) \cdot Y = 0$ is a linear algebra problem.

The intelligence–free approach to linear algebra is to write out the matrix for $(R^{-1} - R)$ and use row operations to reduce it to row–echelon form. I did this. It took a while but I did it. When I got the solutions it was immediately obvious that they could have been found much faster using the rotation argument sketched in

§3.2.3. I was a little annoyed with myself for not having seen the slick argument sooner (I don't really enjoy row operations), but this kind of thing happens all the time in research.

Next I wondered about how area changes as P is moved to $R(P)$ through linear combinations. The first observation was that if the skip polygon has zero area then going along an arc of the circle $r^2+s^2=1$ (in the notation of §3.2.6) doesn't change area. This is the case $k=0$ in §3.2.6. I tried to explore other cases by using an explicit parameterization with polar coordinates: $(r,s)=(\rho(\theta)\cos(\theta),\rho(\theta)\sin(\theta))$. This quickly got complicated, and the parameterizations obscured what was going on. After a while I gave this up and developed the non–parametric version that appears in the text.

The point is that forty years experience as a research mathematician did not prevent me from making mistakes and straying into blind alleys, even in elementary material. The experience did, however, enable me to spot mistakes and blind alleys fairly quickly so I could correct them, or try something different. I believe this illustrates the best goals in learning mathematics:

- *Not* to never make mistakes, but to learn to watch for and recognize them, and then to fix them.
- *Not* to always see what to do, especially if it is clever, but to learn to try things and watch for clues that they could be improved or are unproductive and should be abandoned.

5. Comments for Educators

These are comments for developers and instructors of courses for prospective school teachers. They mainly concern the use of formal definitions, and associated proof strategies.

5.1. **Formal Definitions and Unpacking.** The use of formal definitions, and the unpacking–packing routine described in §1.5 has been standard practice in professional mathematics for over a century. This practice is a compromise between the way people think and the requirements of mathematics:

- As a practical matter, people have to work more-or-less intuitively with conceptual units.
- Mathematical success requires complete reliability.

The challenge, therefore, is to find ways to develop completely reliable intuition. Explicitly and consciously unpacking formal definitions while developing derived properties seems to work pretty well. In fact the effectiveness of this process was probably a key factor in the explosive growth in scope and complexity of mathematics in the last century

The usual routine is: when a definition is introduced, work explicitly with it for a while, typically by deriving secondary properties. After a certain point you should wean yourself (or your students) from the unpacking routine and rely more on intuition and secondary properties. If the intuition is not ready, unpack a little longer.

There are further general comments following an analysis of the projects.

5.2. **The Fraction Project.** The structure of the first project is designed to make the definition–unpacking routine as easy as possible when it is needed, and avoid it when this can be done safely.

5.2.1. *Commutative Rings.* The development takes place in commutative rings. There is a formal definition of these rings and in principle students could go through the formal definition–unpacking routine to develop familiarity with them. However the rules of arithmetic are essentially the same as for integers and real numbers, and students already have fully reliable intuition for these rules. The unpacking routine is not needed, and going through it would increase complexity without any real benefit.

This is why the chapter opens with “There are, of course, axiomatic formulations of these rules (commutative, distributive etc.) but they are already familiar so you can work without thinking about them explicitly.” This is also why this setting was chosen for the development.

5.2.2. *Fractions.* The definition of fractions is introduced in preliminary form in §1.3 and with a subtle problem addressed in §1.9. The preliminary–final division is used to call attention to the role of zero divisors, and provide an opportunity to develop these to the point that they can be worked with intuitively before tackling the subtle version.

The solution-of-equation definition is how fractions are defined in most texts for teachers, see McCrory [3, 4] for a discussion. It is effective in considerably more generality than originally intended, and as a bonus provides a way to manufacture new rings from old (rings of fractions, §1.10). Benefits like this are usually taken to mean that the definition is mathematically “right”.

This definition should be contrasted with the ones proposed for student use, even by mathematicians. The description in Wu [7] is so diffuse it is hard to know where it starts and ends, and even if students get an “understanding” they certainly cannot work with it with any precision. In mathematical terms it is a roadblock rather than a gateway.

Students should continue to unpack the definition of fraction up through the demonstration of the standard facts in §1.9.3. Afterwards they should have internalized the idea accurately enough to skip most of the unpacking in the problems on rings of fractions, §1.10.1.

5.2.3. *Inverses and Zero Divisors.* These concepts are introduced for use in the development of fractions. They are not particularly tricky, and should not be hard to internalize.

Inverses are introduced in §1.4. This is not a totally new concept, but student’s preexisting intuitions may not be sufficiently reliable so we go through the development process anyway. We also want to illustrate the development process itself, and previous familiarity with inverses helps with this. However after identifying inverses in standard rings in §1.6.2, intuition should be developed enough for general use.

Zero divisors are introduced in §1.8. They should be unpacked in Problem 1.9.1. Unpacking may or may not be necessary in the description of zero divisors in standard rings, §1.9.2, and should be unnecessary after that.

5.2.4. *Rings of Fractions.* Fraction rings are introduced to encourage students to expand their understanding of the fraction concept. The original definition is for a *single* fraction. Here the focus is on the *set of all fractions* and the definition becomes a procedure for producing this set. This is a change of perspective more than a new definition.

Fraction rings explain where fractions are defined. A solid answer, in other words, to "what is a fraction?". Unfortunately the answer is "a fraction is an element of the set of all fractions". Answers of this form sound—and usually are—stupid, but here it turns out to be profound.

Sometimes changes of perspective require as much practice and unpacking as genuinely new concepts. This will vary from student to student and can be a difficulty when students are working in groups. Students who "get it" are sometimes impatient with those who don't.

5.2.5. *Ring Homomorphisms.* These are also largely a change of perspective. A full definition is given and students should unpack it when working through the examples. After that, however, they should realize that they have already worked with many examples and should be able to relax about details.

One objective is to connect with something they know. Polynomial fractions, when regarded as functions of a real variable, may not be defined at some values of the variable. This is now seen as an instance of a general phenomenon: a ring homomorphism (evaluation at a number) that does not take a fraction to a fraction because the denominator becomes a zero divisor.

5.2.6. *Grothendieck Groups.* This is again a concept–broadening change of perspective. There are two novelties: working with additive rather than multiplicative notation, and using brute force to fix the zero–divisor problem.

Changing notation (here from $\times$ to $+$) causes serious conceptual dislocation even though it is mathematically inessential. To accommodate this we essentially repeat the development, starting with the definition of commutative semigroup. This should be unpacked while working with the examples in §1.12.2, but this should be enough to establish good contact with previous experience.

The zero–divisor problem for fractions described in §1.7 is revisited in additive notation in §1.12.4, and the alternative fix is described.

Students may vary widely in their need to unpack the definition of Grothendieck group when verifying the identities and working through the examples.

The connection with fractions is nailed down firmly in the last two examples in Problem 1.12.7, and Proposition 1.12.8, by applying the construction to the multiplicative operation in a ring. To avoid notational confusion students should translate the whole Grothendieck definition to multiplicative notation. This should broaden the students' perspective on notation and the nature of mathematical operations as well as the fraction construction.

5.3. **The Area Project.** The main objective of the area project is to show how working without a definition makes a subject more difficult and limits what can be done. Conversely, a definition—or even a good formula—can be quite powerful and can open up rich and unexpected possibilities.

It may be that a definition or effective formula for area is impossible at the school level. In that case it seems particularly important that *teachers* realize that something important is missing.

- If a student has trouble understanding area it may be because the treatment is defective. Teachers should be able to draw on a deeper understanding for explanations, not just repeat something that did not work.
- Small changes in presentation or teacher attitude may help students make the transition to better treatments in later courses.

5.3.1. *Generalized Areas.* The discussion of complex area in §**??**, and "dynamic" area of parameterized families in §**??** have several objectives:

- To suggest that definitions and formulas are not fixed and static, but can be a starting point for exploration.
- To emphasize that there is nothing disgraceful about an exploration that is unsuccessful. It should at least shed light on what makes the successful versions work.
- To re-enforce the idea implicit in the fraction project that flexibility about number systems (rings) can be very useful. Here we see that rings with special properties can be designed to test or enable extensions of a formula.

5.3.2. *Infinitesimals.* As an aside we mention that the ring $\mathbf{R}[\delta]/(\delta^2 = 0)$ used in §**??** is related to old attempts to define derivatives using infinitesimals.

Think of δ as a very small number rather than a formal symbol. The derivative $f'(t)$ is then approximately $\frac{f(t+\delta)-f(t)}{\delta}$. The formula for $f^\delta(t)$ in §**??** then becomes

$$f^\delta(t) = f(t) + \delta f'(t) \simeq f(t) + \delta\,\frac{f(t+\delta)-f(t)}{\delta} = f(t+\delta).$$

This is an attractive heuristic formulation but turns out to be unsatisfactory as a definition. In particular there are severe difficulties getting a precise meaning for "approximately", and as a result the infinitesimal approach has some of the same drawbacks as a heuristic definition of area.

Historically, limits ($\delta \to 0$) were discovered to provide a powerful and flexible definition for the derivative. These replaced infinitesimals in professional mathematics well over a century ago. However taking δ to 0 makes the formula $f(t) + \delta f'(t)$ nonsense.

A way of making precise sense of infinitesimals was finally discovered by Abraham Robinson and others in the 1960s, almost three hundred years after Newton and Leibnitz introduced them as heuristic tools. Robinson's work uses a very sophisticated version of the ring $\mathbf{R}[\delta]/(\delta^2 = 0)$. In principle this means infinitesimal formulas can be used again. However the fine print needed to make it work is so subtle and tricky that this has turned out to be impractical.

The conclusion—again—is that heuristic or intuitive formulations are unsatisfactory for ambitious development and calculation.

5.4. **Cautions about Definitions and Internalization.** Formal definitions are essentially never used in K–12 mathematics. Here we address some justifications given for this, and related potential misunderstandings with these projects.

(1) Näive or innate intuitions are never sufficiently precise. We are able to work with pre–existing intuition about arithmetic in commutative rings in §1.1 because this intuition is not näive: it has been acquired through a great deal of disciplined practice with ordinary numbers and algebra. Problems due to K–12 use of a näive idea of area, see §2.3, are more typical. By comparison with what can be done with a definition, it is hard to compute, hard to describe in alternate ways, and many properties, e.g. invariance under skew transformations, see §2.4.2, remain hidden[12].

[12]A precise definition of area may be impossible in K–12, but realizing there is a problem may open the way for a better partial treatment, see §**??**.

(2) Internalizing properties of an object does *not* render the definition unnecessary, and being able to forget the definition is *not* an objective of the internalization process. In fact a good test of successful internalization is that the student should be able to reproduce a completely precise statement of the definition. This means it is always available for explicit use if there is any doubt, and even experts have to do this from time to time.
(3) Another confusion comes from the perception that experts work with derived properties and intuition, not definitions. This suggests that definitions, unpacking, etc. could be dispensed with, and intuition developed directly from derived properties. Unfortunately this is not the case: reliable intuition is developed by *deriving* the properties, not from the properties themselves. Trying to skip the development process usually leads to dysfunctional understanding that sooner or later will cause trouble, see the comments in §2.3. There is no logical reason for this: it seems to be a feature of human learning. It also seems to be particularly true for less–capable students (i.e. they benefit most from explicit development).
(4) There is a temptation to break concepts into small pieces to make them easier to absorb. However because the pieces have to be assembled after absorption, this actually increases complexity and makes the subject more difficult. Difficult concepts can be approached in stages, c.f. the development of fractions, if the overall conceptual unity is kept clear.
(5) Unpacking definitions is essentially routine and algorithmic. Many students actually enjoy it once they get used to it. However full unpacking is supposed to be a temporary expedient used during the development of reliable intuition, and some students have to be discouraged from continuing when it is no longer appropriate.

5.5. **Summary.**

- Formal definitions provide a repository, training ground, and anchor for intuition.
- Reliable intuitions *incorporate* definitions rather than rendering them unnecessary.
- Good definitions and accompanying development are designed to promote development of reliable intuition.
- Good definitions frequently have unexpected benefits, see the comments at the beginning of §1.12 and the end of §2.3.1.

References

[1] (References to be expanded)
[2] Banchoff, Thomas; *Midpoint Polygons*, 1999 course materials at http://techhouse.brown.edu/ mdp/midpoint/.
[3] McCrory, Raven; *Mathematicians and Mathematics Textbooks for Prospective Elementary Teachers*; Notices of the Amer. Math. Soc. Vol. 53 (2006) pp. 20–29; link http://www.ams.org/notices/200601/fea-mccrory.pdf.
[4] Response to letter about the above; Notices of the Amer. Math. Soc. Vol. 53 (2006) p. 526; link http://www.ams.org/notices/200605/commentary.pdf.
[5] Quinn, Frank; *The Nature of Contemporary Mathematics*, retrieved from http://www.math.vt.edu/people/quinn/education/index.html.
[6] Wilkerson-Jerde, Michelle H.; Wilensky, Uri *Understanding Proof: Tracking Experts Developing Understanding of an Unfamiliar Proof* Proceedings of ICMI Study 19 (2009).

[7] Wu, H. *Fractions, Decimals, and Rational Numbers*, retrieved from http://math.berkeley.edu/ wu/NMPfractions4.pdf.