Math Club Notes: 2005 May 30

1 Triangle decomposition

From the 1982 Putnam:

> Let M be the midpoint of side BC of a general $\triangle ABC$. Using the *smallest possible* n, describe a method for cutting $\triangle AMB$ into n triangles which can be reassembled to form a triangle congruent to $\triangle AMC$.

I leave this fun problem with you.

2 Mean Value Theorem

The Mean Value Theorem (hereafter "MVT") is a fundamental result about differentiable functions in elementary calculus. Here it is:

> If f is continuous on $[a, b]$ and differentiable on (a, b), then there exists a number $c \in (a, b)$ such that
>
> $$f'(c) = \frac{f(b) - f(a)}{b - a} .$$

The RHS is the slope of the secant line connecting $(a, f(a))$ and $(b, f(b))$; the theorem states that f has at least one tangent of the same slope, somewhere between the endpoints of the secant.

As you know, if we think of f as giving the position of some object as a function of time, then the secant slope is the object's average velocity over the interval $[a, b]$, and the tangent slope is its instantaneous velocity at the point of tangency. The theorem states that at some point in the interval, the object actually attains its average velocity. Hence the "mean value" of the name.

(Note that functions that are not differentiable everywhere in (a, b) need not attain their average slope. Consider, for example, $f(x) = |x|$, over the interval $[-1, 1]$.)

As an application of MVT, let's prove this theorem: if f is differentiable on an interval I, then

$$f \text{ is increasing on } I \iff \forall x \in I\colon f'(x) \geq 0 .$$

This is one of a family of theorems which relate a function's derivatives to the shape of its graph.

It is handy to prove this lemma first:

$$f \text{ is increasing on } I \iff \text{all of } f\text{'s secants on } I \text{ have nonnegative slope}$$

A typical definition: we say that f is increasing on an interval I if

$$\forall a, b \in I\colon a < b \implies f(a) \le f(b) .$$

(Note that this is "increasing", not "strictly increasing".) And f's secants on I have nonnegative slope if and only if

$$\forall a, b \in I\colon a \ne b \implies 0 \le \frac{f(b) - f(a)}{b - a} .$$

I leave proving the equivalence of these statements as an exercise.

Now to our theorem. The lemma reduces it to this statement: all f's secants have nonnegative slope if and only if all its tangents have nonnegative slope.

($\Rightarrow$) If all f's secants have nonnegative slope, then for any $x \in I$ we have

$$f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x} \ge 0 ,$$

since the limit of values which are all ≥ 0 must itself be ≥ 0.

($\Leftarrow$) Suppose all f's tangents have nonnegative slope. According to MVT, for any $a, b \in I$ there exists a $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a} .$$

Since $c \in (a, b) \subset I$, by hypothesis $f'(c) \ge 0$; therefore the slope of the secant from a to b is also ≥ 0.

Note how, when proving $\Rightarrow$, we can reason from secant slopes to tangent slopes simply from the definition of the derivative: tangent slopes are limits of secant slopes. When proving $\Leftarrow$, we want to reason from tangent slopes to secant slopes; MVT comes to our aid here, by assuring us that for each secant there is a tangent of the same slope.

Exercise: Formulate and prove the similar theorem for decreasing differentiable functions. (This can be done either by adapting the above proof, or by considering the function $g(x) = -f(x)$.)

Exercise: Suppose f is differentiable on an interval I. The statement

$$f \text{ is strictly increasing on } I \iff \forall x \in I\colon f'(x) > 0$$

is false. Explain how $f(x) = x^3$ is a counterexample. One of the directions is true; which one? Prove it by adapting the proof of that direction given above for "increasing" and "≥ 0". If you try to adapt the above proof of the other direction, what step goes astray?

Exercise: Suppose f is twice-differentiable on an interval I. Prove that the graph of f lies on or above all its tangent lines on I if and only if $f'' \ge 0$ on I. (This is one interpretation of "concave up". The notes for May 22 point out how it can be used to prove Bernoulli's inequality.) Can we say something similar about f's secants?

3 Studying math

What math should one study, and how should one study it?

My feeling is that it almost doesn't matter what math one studies: calculus, algebra, geometry, combinatorics, number theory, it's all good. Ray pointed out (quite astutely, I think) that the courses for (say) the U of A's Honors Math program probably reflect what actual mathematicians think is important. So if you're looking for a field of mathematics to study on your own, pick one that the curriculum writers felt deserved a course.

The how question is more difficult. Some thoughts:

1. When studying a proof, try to identify its major ideas. Often there's just one, and the rest is routine manipulation. A good way to identify these ideas is, ironically, not to read the proof. That is, read the statement of the theorem, then try to prove it yourself. If you succeed, you'll know very well what your idea was. If you don't, when you read the proof you'll spot immediately where the author does something you didn't think of — that step is either the major idea itself, or a necessary maneuver to get to the point of being able to apply the major idea. Study that step.

2. Try to find out what a theorem is for. (Sometimes this is not obvious.) Study how the theorem is used in the proofs of subsequent theorems. Is it just a lemma on the way to something more important? A fundamental result that is used over and over again? A minor result which the author includes just to illustrate how something else can be used? (Bowman's proof of the existence of nth roots, for example, illustrates what IVT is for. He doesn't explain what he's doing, but that's why he gives this example.)

3. Paraphrase. Paraphrase expressions: the quotient $(f(b) - f(a))/(b - a)$ is the slope of a secant. Paraphrase theorems: MVT states that every secant has a parallel tangent. Paraphrase what theorems are for: MVT is for inferring secant slopes from tangent slopes. Paraphrasing like this involves getting a broad conceptual grasp of what theorems are about. It forces you to make your conceptual understanding of things explicit. (Sometimes a paraphrase helps you notice a connection to something else. In the notes for May 22, I paraphrased Euclid I.35 as "shearing preserves area". This reminded me of how determinants in linear algebra express a linear transformation's effect on area, and that led naturally to the rest of the discussion in those notes.)

That's all I can think of at the moment.

4 Combinatorics of distributivity

In the notes for May 22, I used the following manipulation:

$$\sum_i \sum_j a_i b_j = \sum_i \Big(a_i \sum_j b_j \Big) = \Big(\sum_i a_i \Big) \Big(\sum_j b_j \Big)$$

In the first step, we observe that a_i does not depend on j, so we can factor it out of the sum over j. In the second, we observe that $\sum_j b_j$ does not depend on i, so we can factor it out of the sum over i.

This result expresses a fundamental fact about the combinatorics of distributivity. On the right, we have, say,

$$(a_1 + a_2 + a_3)(b_1 + b_2 + b_3 + b_4) \ .$$

On the left we have

$$\begin{aligned} &a_1 b_1 + a_1 b_2 + a_1 b_3 + a_1 b_4 \\ +&a_2 b_1 + a_2 b_2 + a_2 b_3 + a_2 b_4 \\ +&a_3 b_1 + a_3 b_2 + a_3 b_3 + a_3 b_4 \ . \end{aligned}$$

Summing over i and j yields, as shown here, the sum of every possible combination of a_i and b_j. This results from multiplying the sum of all the a_i by the sum of all the b_j.

In a sense, this is why the binomial theorem works. For example,

$$(x+y)^3 = (x+y)(x+y)(x+y) = xxx+xxy+xyx+yxx+xyy+yxy+yyx+yyy \ .$$

Multiplying these three binomials yields a sum of every possible three-factor combination of x and y. When we (use commutativity to) gather like terms, we essentially count how many of those combinations have so many xs and so many ys:

$$\begin{aligned} (x+y)^3 &= xxx + xxy + xyx + yxx + xyy + yxy + yyx + yyy \\ &= (xxx) + (xxy + xyx + yxx) + (xyy + yxy + yyx) + (yyy) \\ &= 1x^3 + 3x^2y + 3xy^2 + 1y^3 \ . \end{aligned}$$

This is why the numbers $\binom{3}{k}$ are both the coefficients in this expansion and the number of ways to choose k objects out of 3. In proofs, one often sees statements such as, "The number of x^2y terms is the number of ways we can get two xs and one y, if we choose either x or y from each of the binomials." This kind of argument is sound because, well, that's how distributivity works.

(Some of the weirder uses of generating functions — e.g., counting tilings of a rectangle — crucially rely on this way of looking at algebraic manipulations from a combinatoric point of view.)