Johns Hopkins Math Tournament 2020

Proof Round: Synthetic Calculus

February 8, 2020

Problem	Points	Score
1	10	
2	5	
3	10	
4	15	
5	10	
6	5	
7	20	
8	10	
9	15	
Total	100	

Instructions

- The exam is worth 100 points; each part's point value is given in brackets next to the part.
- To receive full credit, the presentation must be legible, orderly, clear, and concise.
- If a problem says "list" or "compute," you need not justify your answer. If a problem says "determine," "find," or "show," then you must show your work or explain your reasoning to receive full credit, although such explanations do not have to be lengthy. If a problem says justify or prove, then you must prove your answer rigorously.
- Even if not proved, earlier numbered items may be used in solutions to later numbered items, but not vice versa.
- Pages submitted for credit should be **numbered in consecutive order at the top of each page** in what your team considers to be proper sequential order.
- **Please write on only one side of the answer papers.**
- Put the **team number** (NOT the team name) on the cover sheet used as the first page of the papers submitted. Do not identify the team in any other way.

Introduction

In this proof round, you will approach a somewhat difficult mathematical concept–calculus–from an unconventional angle. Instead of learning calculus as it is usually taught in AP or college courses, you will revisit the very concept of a number and, by learning a thing or two about logic, consider the possibility of the existence of "infinitesimal" numbers. These "infinitesimal" numbers allow all of differential calculus to be developed through simple algebra, provided that certain logical rules underlying usual mathematics are modified. Knowledge of usual calculus is not required, and it will not help you in this proof round.

1 Revisiting Numbers

The number line $\mathcal{R}$ is a straight line on which two points (one called "0" and the other called "1") are marked. The ancient Greeks did arithmetic calculations using a straight edge and a compass with the understanding that a number is a point on this line. The number corresponding to any point is the ratio of its distance to 0 (which may be positive or negative) and the distance between 0 and 1. We will use points on $\mathcal{R}$ and their corresponding number interchangeably from now on.

For your reference, here is a non-exhaustive list of common operations possible using a straight edge and a compass that you may refer to throughout this section.

1. Creating the line through two existing points.
2. Creating the circle through one point with center at another point.
3. Creating the perpendicular bisector and the midpoint of a segment.
4. Drawing a perpendicular line through a point on a given line.
5. Bisecting an angle.
6. Constructing a line parallel to a give line through a given point.

The goal of this short section is to establish that using geometry (straight edge and compass constructions) alone, we can equip $\mathcal{R}$ with many familiar operations.

Example 1.1. *Given two numbers a and b, we may find their sum $a + b$ by using the two legs of the compass to measure out the distance between a and 0, and move the compass so that the leg that used to be at 0 is now at b. The new location of the leg that used to be at a is $a + b$.*

Example 1.2. *Given a number a, $a \neq 0$, we may negate it by constructing a circle of radius a and center 0. The circle intersects the straight line twice. One of these intersections is of course a, and we define the other to be $-a$.*

Problem 1: Show that the following arithmetic operations are possible in $\mathcal{R}$ by describing, as concisely as possible, how to perform them using a straight edge and a compass. At

each step, you may use operations already defined in previous steps or in the examples above. Draw diagrams if necessary. You do not have to justify your response.
(a) (2 points) Subtracting a number from another.
(b) (3 points) Multiplying two numbers.
(c) (3 points) Finding the reciprocal of a nonzero number.
(d) (2 points) Dividing a number by a nonzero number.

(a) To find $a - b$, add $-b$ to a.

(b) To find ab, draw a ray l from 0. Draw a circle of radius b with center 0, and label its unique intersection with l as C. Connect 1 and C, and label this segment m. Construct a line parallel to m through a, and label its unique intersection with l as D. The distance from 0 to D is ab.

(c) To find $1/a$, draw a ray l from 0. Draw a circle of radius 1 with center 0, and label its unique intersection with l as C. Connect a and C, and label this segment m. Construct a line parallel to m through 1, and label its unique intersection with l as D. The distance from 0 to D is $\frac{1}{a}$.

(d) To find $\frac{a}{b}$, multiply a with $\frac{1}{b}$.

A lesser known construction is that of the square root.

Example 1.3. *To take the square root of a, $a > 0$, draw a line segment AB of length $a+1$, with a point P on it so that $\overline{AP} = a$ and $\overline{PB} = 1$. Draw a circle with AB as its diameter, and draw a ray from P that is perpendicular to AB. Label the unique intersection of P with the circle as D. Then it is claimed that $\overline{PD} = \sqrt{a}$.*

Problem 2: (5 points) Prove that in the construction described above, $\overline{PD} = \sqrt{a}$.

Connect AD and DB. Since $\angle ADB = 90°$, $\angle ADP = \angle DBP$ and $\angle DAP = \angle PDB$. Therefore $\triangle APD \sim \triangle DPB$. As a result, $\frac{\overline{AP}}{\overline{PD}} = \frac{\overline{PD}}{\overline{PB}}$, hence $\overline{PD} = \sqrt{a}$.

We've now established the following:

1. There is an operation called addition on $\mathcal{R}$. This operation is associative.
2. There is a point 0 such that for any number a, $a + 0 = a$.
3. For every number a, there exists an additive inverse $-a$ such that $a + (-a) = 0$.
4. There is an operation called multiplication on $\mathcal{R}$. This operation is associative and commutative.
5. There is a point 1 such that for any number a, $1 \times a = a$.
6. If $a \neq 0$, there exists a multiplicative inverse $\frac{1}{a}$ such that $a \times \frac{1}{a} = 1$.
7. The distributive law for addition and multiplication holds.

Definition 1.1. A structure satisfying this list of properties is called a field. The list of properties is called the field axioms.

You did not prove the final statement geometrically, but it should not be a surprise that it holds. The rest of the proof round will introduce the infinitesimal numbers by "abusing a loophole" in the sixth statement above. Of course, the real numbers (or the rational numbers) satisfy all of the above axioms. But is it not possible to introduce a class of points so close to 0 that they are not nonzero (hence don't have multiplicative inverses), yet not necessarily zero? In other words, can we change logic just a little bit so that a number does not need to be either zero or nonzero?

2 The Infinitesimal Numbers

The definition of the infinitesimal numbers is straightforward.

Definition 2.1. The collection of infinitesimal numbers $\mathcal{D} = \{d \in \mathcal{R} | d^2 = 0\}$.

At first sight, this set seems trivial.

Proposition 2.1. $\mathcal{D} = \{0\}$.

Proof. Consider $d \in \mathcal{D}$. If $d = 0$, we are done. If $d \neq 0$, then by the field axioms, d has a multiplicative inverse $\frac{1}{d}$. Multiplying both sides of $d^2 = 0$ by $\frac{1}{d}$ we get $d = 0$, which is absurd. Hence $\mathcal{D} = \{0\}$. □

But for reasons that will become clear later, we would like this collection to be nontrivial. We want it to contain all the numbers that are so close to zero they are indistinguishable from it. To do this, let us observe that the proof above uses an assumption that is not part of the field axioms.

Definition 2.2. The Principle of Excluded Middle (PEM) states that a proposition is either true or false.

In the previous proof, we assumed that d is either zero or nonzero. This allowed us to consider the two cases and conclude that of the numbers that are either zero or nonzero, only zero is in the set $\mathcal{D}$.

However, proofs that use PEM do not always make sense. For example, consider the following.

Proposition 2.2. *There are two irrational numbers α and β such that α^β is rational.*

Problem 3: Using PEM, prove the proposition by considering the number $\sqrt{2}^{\sqrt{2}}$.

- **(a)** (5 points) Assume that $\sqrt{2}^{\sqrt{2}}$ is rational, show that the proposition holds.
- **(b)** (5 points) Assume that $\sqrt{2}^{\sqrt{2}}$ is irrational, show that the proposition holds.

a $\alpha = \beta = \sqrt{2}$

b Let $\alpha = \sqrt{2}^{\sqrt{2}}$, $\beta = \sqrt{2}$, then $\alpha^\beta = \sqrt{2}^2 = 2$, which is rational.

This proof is deeply unsatisfying: we never actually figure out what the α and β in the proposition are!

The validity of PEM (whether it is as intuitive as other logical laws) is, in some way, still an active debate within the mathematics community. We know that there will always be propositions that are neither provably true nor provably false (the Continuum Hypothesis, for example). Without getting into this debate's history and current situation, let us entertain a mathematical world in which PEM does not hold. What would it look like?

The infinitesimal numbers introduced by the following axiom provides a possible answer.

Axiom 2.1. *Every function $f : \mathcal{D} \to \mathcal{R}$ is of the form*

$$f(d) = a + bd$$

For some unique $a, b \in \mathcal{R}$

Let us assume that this axiom holds in our $\mathcal{R}$, and let us begin exploring some of its consequences. For the rest of this proof round, PEM will not longer be valid.

Problem 4: Prove the following using the definition of infinitesimals, the field axioms, and Axiom 2.1.
(a) (5 points) Prove that if $ad = bd$ for all $d \in \mathcal{D}$, then $a = b$.
(b) (5 points) Prove that $\mathcal{D} \neq \{0\}$. (hint: assume $\mathcal{D} = \{0\}$ and consider two functions from $\mathcal{D}$ to $\mathcal{R}$ with the same a term but different b terms, as in the axiom. Are they different functions?)
(c) (5 points) Prove that the products of infinitesimals are infinitesimals (that is to say, their squares are 0).

a Consider $f(d) = ad = bd$ as a function from $\mathcal{D}$ to $\mathcal{R}$. Since $ad = bd$ for all $d \in \mathcal{D}$, this is a well defined function. But by the axiom, the coefficient of d is necessarily unique, hence $a = b$.

b Consider two functions $f(d) = ad$ and $g(d) = bd$, $a \neq b$, both as functions from $\mathcal{D}$ to $\mathcal{R}$. If $\mathcal{D} = \{0\}$, these are the same function, hence by the axiom $a = b$, which is absurd.

c Let $d_1, d_2 \in \mathcal{D}$. Then $(d_1 d_2)^2 = d_1^2 \times d_2^2 = 0$, hence their product is in $\mathcal{D}$.

3 Differential Calculus

Let us consider the slope of a function. For a linear function (a function whose graph is a straight line), the slope can be described as the change of output divided by change of input. For example, to find the slope of $y = 2x$, we may fix a point (say $x = 0$). Then for each unit of change in x (say from 0 to 1), we can observe y is changed by twice that amount (from 0 to 2). We can hence conclude that the slope of this function is 2.

What happens if the function we describe is not linear? First, the fixed point we choose is not longer arbitrary: the slope of the function may be different at different points of the function. The function $y = 1/x$, for example, is very flat when x is large and very steep when x close to 0.

Second, changing the input by some nontrivial value (1 in the above example) might no longer be optimal. If we try to find the slope of $y = x^2$ by changing x from 0 to 1 we will arrive at the conclusion that the parabola has slope 1 at the origin, which is absurd: we can see it's flat! To find the slope of a curve in general, then, we would like to change the input by a very very small amount–an infinitesimal amount, in fact–and observe the change in output proportional to the change of input.

To compute the slope of arbitrary functions as a function: this is the goal of differential calculus.

Definition 3.1. Given a function $f(x) : \mathcal{R} \to \mathcal{R}$, we may fix some x in R and consider a function from $\mathcal{D}$ to $\mathcal{R}$ defined as

$$f_x(d) = f(x + d)$$

By Axiom 2.1, there is an unique constant c_x such that

$$f_x(d) = f(x) + c_x \cdot d$$

This c_x is the derivative (that is to say, slope) of $f(x)$ at this x.

We can now find the slope of nonlinear functions at specific points. For example,

Proposition 3.1. *The derivative of x^2 at $x = 2$ is 4.*

Proof. Here $f(x) = x^2$. Since $x = 2$, $f_2(d) = f(2 + d) = (2 + d)^2 = 4 + 4d + d^2 = 4 + 4d$ since $d^2 = 0$. Thus $f'(2) = c_2 = 4$. $\square$

Problem 5: Using the definition given above, find the derivative of the following functions at the given points. Show your work.

(a) (5 points) $f(x) = 8, x = 5$.

(b) (5 points) $f(x) = \frac{1}{x}, x = 2$.

a $f_5(d) = 8 + 0 \cdot d$. $f'(5) = 0$.

b $f_2(d) = \frac{1}{2+d} = \frac{2-d}{4-d^2} = \frac{1}{2} - \frac{d}{4}$. $f_2(d) = -\frac{1}{4}$.

Furthermore, as we vary the point x at which we compute the derivative, we may express the slope of a function as a function in itself.

Definition 3.2. The *derivative function* of $f(x)$ is defined as

$$f'(x) = c_x$$

In other words, we define $f'(x)$ to be the (unique) function such that

$$f(x + d) = f(x) + d \cdot f'(x)$$

The derivative function therefore takes a x value and finds the ratio of infinitesimal change of output and input at that x. The derivative is always defined here because what Axiom 2.1 says, in effect, is that every function is "locally"–i.e., plus or minus some infinitesimal number–a straight line.

For example, since for $f(x) = x^2$ and for each x,

$$f(x+d) = (x+d)^2 = x^2 + 2dx + d^2 = x^2 + 2dx$$

We get, from this definition, that $(x^2)' = 2x$. That is to say, for every x, the slope of the parabola x^2 at x is always $2x$. The slope of the parabola at $x = 1$ is 2, the slope at $x = 10$ is 20, and so on.

Problem 6: (5 points) Using the definition given above, find the derivative function of $x^3 - x + 1$.

$f(x+d) = (x^3 + 3x^2 d + 3xd^2 + d^3) - (x+d) + 1 = (x^3 - x + 1) + d \cdot (3x^2 - 1)$. Hence $f' = 3x^2 - 1$.

Problem 7: Let f, g be functions from $\mathcal{R}$ to $\mathcal{R}$. Prove the following rules of differentiation.

(a) (2 points) If c is a constant, $(c \cdot f)' = c \cdot f'$.
(b) (3 points) $(f+g)' = f' + g'$.
(c) (5 points) $(f \cdot g)' = f \cdot g' + f' \cdot g$.
(d) (5 points) $(f/g)' = (g \cdot f' - f \cdot g')/g^2$.
(e) (5 points) $(f \circ g)' = f'(g) \cdot g'$.

a $c \cdot f(x+d) = c \cdot f(x) + c \cdot d \cdot f'(x)$.

b $(f+g)(x+d) = (f+g)(x) + d \cdot (f+g)'(x) = f(x) + d \cdot f'(x) + g(x) + d \cdot g'(x)$.

c $(f \cdot g)(x+d) = f(x+d) \cdot g(x+d) = (f(x) + df'(x)) \cdot (g(x) + dg'(x)) = (f \cdot g)(x) + d(g(x) \cdot f'(x) + f(x) \cdot g'(x)) + d^2(f'(x)g'(x)) = (f \cdot g)(x) + d(g(x) \cdot f'(x) + f(x) \cdot g'(x))$.

d $(f/g)(x+d) = f(x+d)/g(x+d) = (f(x) + df'(x))/(g(x) + dg'(x)) = (f + df')(g - dg')/(g^2 - d^2(g')^2) = (fg + d(gf' - fg'))/g^2 = f/g + d((gf' - fg')/g^2)$.

e $(f \circ g)(x+d) = f(g(x) + dg'(x)) = f(g)(x) + d \cdot g'(x) \cdot f'(g(x))$. The second step is possible because $d \cdot g'(x) \in \mathcal{D}$.

4 Kepler's Third law of Motion

As we have hinted, the implication of Axiom 2.1 is that every function is locally a straight line. We will now explore this geometric interpretation of Axiom 2.1 by proving Kepler's Third Law of Motion, which states that the area spanned by the cord connecting any planetary body to the body it orbits around is a constant function of time.

We begin by finding the derivative of sine and cosine. We assume the usual trigonometric identities for all of $\mathcal{R}$.

1. $\sin 0 = 0$
2. $\cos 0 = 1$
3. $\sin^2 x + \cos^2 x = 1$
4. $\sin(x + y) = \sin x \cos y + \cos x \sin y$
5. $\cos(x + y) = \cos x \cos y - \sin x \sin y$

To find the derivative of sine, consider a circle of unit radius and a segment AB of the circle of angle $2d$, where $d \in \mathcal{D}$. By the geometric interpretation of Axiom 2.1, since the arc AB is of length $2d$, we may in fact consider it a straight segment, hence coinciding with the cord AB. Since the cord AB has length $2 \sin d$, we conclude that $\sin d = d$ for all $d \in \mathcal{D}$.

Problem 8:
(a) (2 points) Use trigonometric identities and $\sin d = d$ to conclude that $\cos d = 1$ for all $d \in \mathcal{D}$.
(b) (4 points) Use trigonometric identities and the equations proven above to find the derivative function of sine.
(c) (4 points) Similarly, find the derivative function of cosine.

a Since $\sin^2 d + \cos^2 d = 1$ and $\sin^2 d = d^2 = 0$, we have $\cos d = 1$.

b $\sin(x + d) = \sin x \cos d + \cos x \sin d = \sin x + d \cos x$, hence $(\sin x)' = \cos x$.

c $\cos(x + d) = \cos x \cos d - \sin x \sin d = \cos x - d \sin x$, hence $(\cos x)' = -\sin x$.

We are now ready to tackle Kepler's Third Law. Consider a particle moving under the influence of the gravitation force of particle O, and write P for its current position (at time $t = t_0$). Construct a Cartesian coordinate system with O as its origin, describing P using the usual tuple x, y. We use r to denote the length of the cord OP and θ the angle formed between this cord and the positive ray of the x-axis. All four measurements x, y, r, and θ are functions of t, and the usual conversion formulae stand between the Cartesian and the polar measurements:

$$x(t) = r(t) \cos \theta(t)$$
$$y(t) = r(t) \sin \theta(t)$$

As this is not a physics competition, it suffices for us to remark that, if a body of great mass (say, our Sun) occupies O and a body of much smaller mass (say, our Earth) occupies P, then the Newtonian theory of gravity guarantees that the following formula holds for the position of the moving particle:

$$xy'' - yx'' = 0$$

Here, x'' denotes the double derivative (that is to say, the derivative function of the derivative function) of the function x with respect to t, and the same holds for y''.

Problem 9:

(a) (3 points) Suppose, after a short period of time d, where $d \in \mathcal{D}$, the new position of the particle that was at P is now Q. Using our axiom and basic geometry, show that the area spanned by the cord connecting O and the moving particle during this short period of time is $\frac{1}{2}dr^2(t_0)\theta'(t_0)$.
(b) (3 points) Show that $A'(t_0) = \frac{1}{2}r^2(t_0)\theta'(t_0)$, where $A(t)$ is the area spanned by the cord connecting O and the moving particle.
(c) (3 points) Compute x' and y' in terms of θ and r using the coordinate conversion equations above.
(d) (6 points) Show that $A'' = 0$. Since the area function has no acceleration with respect to time (that is to say, its speed has a zero derivative function), it must be a linear function, therefore we have proven Kepler's Third Law of Motion!

a Using the axiom, we may assume that the trajectory of the particle from P and Q is straight. Then the area spanned by the cord is a triangle $\triangle OPQ$. The base is $r(t_0 + d) = r(t_0) + dr'(t_0)$. The height is $r(t_0)\sin(d\theta'(t_0)) = r(t_0)d\theta'(t_0)$. Hence the area of $\triangle OPQ$ is $\frac{1}{2}(r(t_0) + r'(t))(r(t_0)d\theta'(t_0)) = \frac{1}{2}r^2(t_0)d\theta'(t_0)$.

b Since $A(t_0 + d) - A(t_0)$ is the area of $\triangle OPQ$, and $A(t_0 + d) - A(t_0) = dA'(t_0)$, we conclude that $\frac{1}{2}r^2(t_0)\theta'(t_0) = dA'(t_0)$. Since this equation clearly holds for all $d \in \mathcal{D}$, we use an earlier result to conclude $A'(t_0) = \frac{1}{2}r^2(t_0)\theta'(t_0)$.

c

$$x' = r'\cos\theta - r\theta'\sin\theta$$

$$y' = r'\sin\theta + r\theta'\cos\theta$$

d Observe that $xy' - x'y = r^2\theta' = 2A'$. Taking the derivative on both sides yields

$$2A'' = x'y' + xy'' - y'x' - yx'' = xy'' - yx'' = 0$$

.