Hyperreal Numbers: An Elementary Inquiry-Based Introduction

Handouts for a course from Canada/USA Mathcamp 2017

Don Laackman
Problem 1. At Mathcamp, in the days when inner tubes were used for many games, from inner tube water polo, to filling innocent staff members’ rooms, it was important for the inner tubes to be filled up by a collection of staff. The campers would send staff off to fill up inner tubes; campers labeled sets $Y$ of staff as “negligible” if no inner tubes got filled, and “substantial” if some inner tubes got filled.

The campers noticed that if sets $Y_1$ and $Y_2$ of staff are “negligible,” then so is their union $Y_1 \cup Y_2$, and a set $Y$ of staff is “negligible” precisely when the rest of the staff form a “substantial” set.

Show that there is a single staff member whose contributions to filling inner tubes is “substantial” and that the total contribution of all the other staff is “negligible.”
A new lounge has been discovered! The campers need to vote on whether it should become the “Rose Lounge,” the “Goldenrod Lounge,” or the “Viridian Lounge.” Lara will be counting the votes, and she promises two things:

1. If all campers vote for the same color, that’s what we’ll call the lounge.
2. If one collection of votes would lead to a particular color for the lounge, and every camper changes their vote, the outcome of the election will also change.

**Problem 2.** For most of this problem, we’ll arbitrarily break the campers up into two subsets: the “hungry campers” and the “other campers.” Every camper is in exactly one of those two sets.

1. Suppose all of the hungry campers vote for Goldenrod, and all of the other campers vote for Rose. What colors could win the election?

2. Let’s assume for now that when hungry campers all vote for Goldenrod and other campers all vote for Rose, Goldenrod wins the election. If all of the hungry campers change their vote to Viridian, and all of the other campers change their votes to some mix of Goldenrod and Viridian (an assignment of votes we’ll call “situation midnight”), what color will win the election?

3. Now suppose the hungry campers switch back to Goldenrod, and all of the other campers switch between Viridian and Goldenrod. What color must now win the election?
(4) Suddenly, all the hungry campers vote for Rose, and all the other campers return to their votes from “situation midnight.” What color wins this election?

(5) Call “situation daylight” a collection of votes from the other campers that differs from every vote in “situation midnight” (so each vote is either for Rose, or the color of Goldenrod and Viridian that that camper doesn’t vote for in “situation midnight”). Show that if all of the hungry campers vote for color X, and the other campers are in “situation daylight,” color X must win the election (hint: compare to each of the scenarios with hungry campers voting for a different color, and the other campers in “situation midnight.”) Since “situation midnight” was arbitrary, so is “situation daylight” - what does this mean for the outcome of the election if all of the hungry campers vote the same way?

(6) If all the hungry voters vote for a mix of 2 colors, X and Y (where X and Y are some distinct pair from \{Rose, Goldenrod, Viridian\}), show that one of X and Y must win the election.

(7) Now consider a - “situation room” - where the hungry campers have some arbitrary mix of votes. Assume that there are two ways for the other campers to vote - “situation normal” and “situation weird” - such that if hungry campers vote as in “situation room” and other campers vote as in “situation normal” a different color wins than if hungry campers vote as in “situation room” and other campers vote as in “situation weird.” Call the colors that win in those two scenarios X and Y; have every hungry camper both change their vote from “situation room” and vote for one of X and Y, and have the other campers vote for a color they don’t vote for in either “situation normal” or “situation weird.” Show that this leads to a contradiction in who must win the election.
(8) Explain why any way of breaking the campers into two disjoint subsets will have exactly one of the subsets choosing the outcome of the election (remember, our assignment of “hungry” and “other” was arbitrary).

(9) Call any subset of campers whose votes decide the election “big.” The complement of a “big” set of campers will be “little,” and if you change any number of votes strictly within a little set, you can’t change the election results. Explain why any set of campers containing a big set is big, and the finite intersection of big sets of campers is big.

(10) Show that there is a single camper whose vote decides the entire election.
Definition. A Filter on some set $X$ is a collection $\mathcal{F}$ of subsets of $X$, such that:

1. $X \in \mathcal{F}$, $\emptyset \notin \mathcal{F}$
2. If $Y_1 \in \mathcal{F}$ and $Y_1 \subseteq Y_2$, then $Y_2 \in \mathcal{F}$
3. If $Y_1, Y_2 \in \mathcal{F}$, then $Y_1 \cap Y_2 \in \mathcal{F}$

A filter $\mathcal{F}$ such that $Y \in \mathcal{F}$ if and only if $X - Y \notin \mathcal{F}$ is called an Ultrafilter on $X$.

Problem 3. 
(1) Suppose $X$ is a finite set; explain why for any ultrafilter $\mathcal{F}$ on $X$, there is a single $a \in X$ such that $Y \in \mathcal{F}$ if and only if $a \in Y$. Such an ultrafilter is called principal.
(2) If an ultrafilter were going to be nonprincipal, what sets must automatically be contained in the ultrafilter?
(3) Start with the filter $\mathcal{F}$ just of sets given by your answer to the previous part; suppose $Y$ and its complement $X - Y$ are both not in $\mathcal{F}$. Show that adding $Y$, and all of its intersections with the sets in $\mathcal{F}$, and all sets containing those intersections, to $\mathcal{F}$ still leaves you with a filter.
(4) Prove that this filter can thus be extended to an ultrafilter, by choosing sets outside of it over and over.
Problem 4. Suppose we fix some nonprincipal ultrafilter $\mathcal{F}$ on $\mathbb{N}$. Show that there is a function, $\lim^*$, taking any bounded sequence of real numbers to a number, such that:

- If $\{x_n\}$ is convergent, then $\lim^* x_n = \lim x_n$.
- $\lim^*(x_n + y_n) = \lim^* x_n + \lim^* y_n$.
- $\lim^*(c \cdot x_n) = c \cdot \lim^* x_n$.
- $|\lim^* x_n| \leq A$ if $|x_m| \leq A$ for all $m$. 

Definition. A nonzero number $\epsilon$ is called infinitesimal if for all $n = 1, 2, 3, \ldots$,

$$|\epsilon| < \frac{1}{n}$$

When this is the case, its reciprocal, $\omega = \frac{1}{\epsilon}$, will be called infinitely large, because for all $n = 1, 2, 3, \ldots$

$$|\omega| > n$$

Problem 1. • Are there any infinitesimal rational numbers?

• Are there any infinitely large rational numbers?

• Are there any infinitesimal real numbers? (You can think of real numbers as decimal sequences.)

• Are there any infinitely large real numbers?
Definition. $\mathbb{R}^N$ is the set of all infinite sequences of real numbers, $\{(a_0, a_1, a_2, \ldots) | \forall i \in \mathbb{N}, a_i \in \mathbb{R}\}$.

Problem 2. (1) How would you define addition and multiplication on $\mathbb{R}^N$?

(2) What are the additive and multiplicative identities for these operations?

(3) Is there a copy of $\mathbb{R}$ as a subset of $\mathbb{R}^N$, using the same addition and multiplication from the previous part?

(4) Can you divide by any sequence that isn’t $(0, 0, 0, \ldots)$?

(5) Given two sequences $(a_i)$ and $(b_i)$, what would it mean to say $(a_i) \leq (b_i)$?
Let's try to make sure that our idea of indexes voting on the properties of a number won't run into any problems. At a minimum, we certainly need “the election says these two numbers are the same” to be an equivalence relation.

**Definition.** A binary relation $\sim$ on a set $X$ is an equivalence relation if, for all $a, b, c \in X$

1. $a \sim a$
2. If $a \sim b$, then $b \sim a$
3. If $a \sim b$ and $b \sim c$, then $a \sim c$

Let’s denote our “an election says these two sequences are the same” relation by $\equiv$

**Problem 3.**

1. What does transitivity (If $(a_i) \equiv (b_i)$ and $(b_i) \equiv (c_i)$, then $(a_i) \equiv (c_i)$) tell you about the winning coalitions in our elections?

2. Define $(r_i) \not\equiv (s_i)$ analogously to $(r_i) \equiv (s_i)$. What must be true of the set of winning coalitions if, for every pair of sequences $(r_i), (s_i)$, either $(r_i) \equiv (s_i)$ or $(r_i) \not\equiv (s_i)$?

3. Using the requirements deduced above, suppose for the moment that there were a finite collection of indices that could decide the election. What would this tell us? (Think about yesterday!)
Definition. A Filter on some set $X$ is a collection $\mathcal{F}$ of subsets of $X$, such that:

1. $X \in \mathcal{F}, \emptyset \notin \mathcal{F}$
2. If $Y_1 \in \mathcal{F}$ and $Y_1 \subseteq Y_2$, then $Y_2 \in \mathcal{F}$
3. If $Y_1, Y_2 \in \mathcal{F}$, then $Y_1 \cap Y_2 \in \mathcal{F}$

A filter $\mathcal{F}$ such that $Y \in \mathcal{F}$ if and only if $X - Y \notin \mathcal{F}$ is called an Ultrafilter on $X$.

Problem 4. (1) Suppose $X$ is a finite set; explain why for any ultrafilter $\mathcal{F}$ on $X$, there is a single $a \in X$ such that $Y \in \mathcal{F}$ if and only if $a \in Y$ (we’ve already proven this). Such an ultrafilter is called principal.

(2) If an ultrafilter on an infinite set (say, $\mathbb{N}$) were going to be nonprincipal, what sets must automatically be contained in the ultrafilter?

(3) Start with the filter $\mathcal{F}$ just of sets given by your answer to the previous part; suppose $Y$ and its complement $X - Y$ are both not in $\mathcal{F}$. Show that adding $Y$, and all of its intersections with the sets in $\mathcal{F}$, and all sets containing those intersections, to $\mathcal{F}$ still leaves you with a filter.

(4) Prove that this filter can thus be extended to an ultrafilter, by choosing sets outside of it over and over.
Let’s finally define the set we’ve been building towards.

**Definition.** The Hyperreal Numbers, $^\ast\mathbb{R}$, is the set of equivalence classes of sequences of real numbers, where two sequences $(r_1, r_2, r_3, \ldots)$ and $(s_1, s_2, s_3, \ldots)$ are the same if an “election” of the indices claims they are the same. Elections are determined by using a nonprincipal ultrafilter on the indices, which in particular, cannot contain any finite set of indices.

**Problem 1.**

1. What hyperreal number is $(0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, \ldots)$, where the $i^{th}$ term is 1 if and only if $i$ is prime, and 0 otherwise?

2. Suppose a hyperreal number can be represented as $(a_0, a_1, a_2, \ldots)$ with all of the $a_i$ elements of some finite set $X$. What can you say about it?

3. Show that any hyperreal can be represented by a sequence that is entirely positive, entirely negative, or entirely 0.

4. Challenge: can every hyperreal number be represented by a sequence that is either entirely increasing, entirely decreasing, or constant?
Definition. Suppose \( b \) is a hyperreal number.

- \( b \) is infinitesimal if, for all \( n \in \mathbb{N} \), \( |b| < \left( \frac{1}{n}, \frac{1}{n}, \frac{1}{n}, \ldots \right) \)
- \( b \) is finite if, for some \( n \in \mathbb{N} \), \( |b| < (n, n, n, \ldots) \)
- \( b \) is infinitely close to the hyperreal \( c \), denoted \( b \simeq c \), if \( b - c \) is infinitesimal. The halo of \( b \) is the set of hyperreals infinitely close to \( b \):
  \[
  \text{hal}(b) = \{ c \in \mathbb{R}^* | b \simeq c \}
  \]
- \( b \) is a finite distance from the hyperreal \( c \), denoted \( b \sim c \), if \( b - c \) is finite. The galaxy of \( b \) is the set of hyperreals of finite distance from \( b \):
  \[
  \text{gal}(b) = \{ c \in \mathbb{R}^* | b \sim c \}
  \]

Problem 2. Show the following

1. \( \simeq \) and \( \sim \) are equivalence relations.

2. \( \text{hal}(b) = \{ b + \epsilon | \epsilon \in \text{hal}(0) \} \)

3. \( \text{gal}(b) = \{ b + c | c \in \text{gal}(0) \} \)

4. Any galaxy contains members of \( \mathbb{Z}^*, \mathbb{Q}^* - \mathbb{Z}^* \), and \( \mathbb{R}^* - \mathbb{Q}^* \) (hyperintegers, hyperfractions, and hyperirrationals).
Problem 3.  

(1) Show that any finite hyperreal number has a unique real number in its halo - this real number is called the “shadow” of the hyperreal.

(2) Show that every hyperreal is infinitely close to some hyperrational.

(3) Draw a picture of the hyperreals based on the standard number line, plus what you’ve shown about the hyperreals.

The problem above show that, if you were of a particularly determined mindset, you could actually define the real numbers by first taking the finite hyperrationals, and then defining the reals as the set of halos. This seems like a lot of extra machinery, but really, it’s pretty close to our usual methods.
Definition. Given any function $f : \mathbb{R} \to \mathbb{R}$, we can hyper-extend $f$ to be defined on the hyperreals by just saying $f^*(s_1, s_2, s_3, \ldots) = (f(s_1), f(s_2), f(s_3), \ldots)$.

A function $f : \mathbb{R} \to \mathbb{R}$ is defined to be continuous at a real number $c$ if $f^*(x) \simeq f^*(c)$ whenever $x \simeq c$.

Problem 4. (1) Is our definition of hyper-extending functions well defined (does it not depend on the choice of representative sequence)?

(2) Use the above definition, the sum formula for sines, and the fact that for any $\theta$, $|\sin(\theta)| \leq |\theta|$, prove that sine is continuous at all real numbers.

(3) Prove that any polynomial function is continuous at every real number.

(4) Prove that

$$f(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

is not continuous at $0$. 
Definition. If \( f \) is defined at \( x \in \mathbb{R} \), then the real number \( L \in \mathbb{R} \) is defined to be the derivative of \( f \) at \( x \) if, for every infinitesimal \( \epsilon \), \( f(x + \epsilon) \) is also defined, and

\[
\frac{\epsilon f(x + \epsilon) - f(x)}{\epsilon} \approx L
\]

If you’ve seen derivatives before, this should look familiar; it’s what the normal definition wishes it could be, giving the instantaneous slope of the function \( f \) at \( x \).

Problem 5. (1) If \( f \) is differentiable at \( x \), show that \( f \) must also be continuous at \( x \).

(2) Use our definition to directly compute the derivatives of \( x^n \), \( \sin(x) \), and \( e^x \) (you can assume that \( \frac{e^\epsilon - 1}{\epsilon} \approx 1 \) for any infinitesimal \( \epsilon \); this is one way to define \( e \)).
Definition. Hyper-Extension

Given any function $f : \mathbb{R} \to \mathbb{R}$, we can hyper-extend $f$ to a function $f^*$ defined on the hyperreals by saying $f^*((s_0, s_1, s_2, \ldots)) = (f(s_0), f(s_1), f(s_2), \ldots)$.

Given a set $A \subseteq \mathbb{R}$, we can we can hyper-extend $A$ to a set $^*A \subseteq ^*\mathbb{R}$ by saying $^*A = \{(a_1, a_2, a_3, \ldots) | a_i \in A\}$.

Problem 1. (1) What did we prove yesterday about the hyper-extension of a finite set $A$?

(2) Give an example of a function $g : ^*\mathbb{R} \to ^*\mathbb{R}$ that is not the hyper-extension of any function $f : \mathbb{R} \to \mathbb{R}$.

(3) Give an example of a set $B \subseteq ^*\mathbb{R}$ that is not the hyper-extension of any set $A \subseteq \mathbb{R}$.

(4) Think of a few properties of functions and sets. Which of these properties are preserved by hyper-extension? Which properties aren’t preserved, but if you held a vote of the indices, they’d tell you that it was?
**Theorem.** The Transfer Principle

Given any first-order logical statement about $\mathbb{R}$ — a statement built out elementary arithmetic statements, “and,” “or,” “not,” and finitely many instances of “$\forall$” and “$\exists$”, such that each quantifier is only quantifying elements of a set, not sets themselves — its truth is equivalent to the statement over $\ast\mathbb{R}$ which is the same, except having every set and function hyper-extended.

The transfer principle holds, in principle, because elementary arithmetic statements about hyperreals are decided by a vote of the indices, and so any statement that is true about the real numbers must be true about the hyperreals. It gets a little trickier when we add in logical connectives and quantifiers - the latter requiring us to hyper-extend things in order to keep the truth of our statements - but it still works out. We won’t be proving this theorem in class, but if you’d like to prove it, come talk to us during TAU.

**Problem 2.** Here is a disproof of the Transfer Principle: The statement “$\forall x \in \mathbb{R}$, there is some natural number $n$ such that $|x| < \frac{1}{n}$” is a true, first-order statement. The equivalent statement over the hyperreals is “$\forall x \in \ast\mathbb{R}$, there is some natural number $n$ such that $|x| < \frac{1}{n}$.” Yet the latter statement is contradicted by the existence of infinitesimals. Thus, transfer must not hold.

What’s wrong with this proof?

**Problem 3.** Prove the infinitude of primes, using the set of primes $P$, the hypernatural number $N! + 1 = (0! + 1, 1! + 1, 2! + 1, 3! + 1, \ldots)$, and transfer.
Recall from yesterday:

**Definition.** A function \( f : \mathbb{R} \to \mathbb{R} \) is defined to be continuous at a real number \( c \) if, for any \( x \in \mathbb{H} \), \( f^*(x) \approx f^*(c) \) whenever \( x \approx c \).

**Problem 4.** Does this make sense as a definition of continuity? Explain why or why not.

**Problem 5.** Prove the Extreme Value Theorem — given a function \( f : \mathbb{R} \to \mathbb{R} \), that is bounded and continuous on some closed interval \([a,b]\), it attains a maximum and minimum value on the interval. (Hint: try spacing out \( n \) points through the interval; what can you say about \( f \) achieving a max and a min just for these points? Then, apply transfer.)

**Problem 6.** Prove the Intermediate Value Theorem - given a function \( f : \mathbb{R} \to \mathbb{R} \) that is continuous on the interval \([a,b]\), and given any number \( c \) such that \( f(a) \leq c \leq f(b) \) or \( f(b) \leq c \leq f(a) \), there is some \( x \in [a,b] \) such that \( f(x) = c \). (Hint: think of how you proved the last problem.)
Definition. If $f$ is defined at $x \in \mathbb{R}$, then the real number $L \in \mathbb{R}$ is defined to be the derivative of $f$ at $x$ if, for every infinitesimal $\epsilon$, $f(x + \epsilon)$ is also defined, and

$$\frac{f(x + \epsilon) - f(x)}{\epsilon} \approx L$$

If you’ve seen derivatives before, this should look familiar; it’s what the normal definition wishes it could be, giving the instantaneous slope of the function $f$ at $x$.

Problem 7. (1) If $f$ is differentiable at $x$, show that $f$ must also be continuous at $x$.

(2) Use our definition to directly compute the derivatives of $x^n$, $\sin(x)$, and $e^x$ (you can assume that $e^{\epsilon - 1} \approx 1$ for any infinitesimal $\epsilon$; this is one way to define $e$).

Problem 8. Big-O notation is used to say that one function eventually grows no faster than another; it is formally defined, for functions $f, g : \mathbb{R} \to \mathbb{R}$, by saying that $f \in O(g)$ if there are some $x_0 \in \mathbb{R}$ and $M \in \mathbb{R}$ such that, for any $x > x_0$, $|f(x)| \leq M|g(x)|$.

Little-o notation is similar, just saying that a function eventually grows much slower than another; it is formally defined, for functions $f, g : \mathbb{R} \to \mathbb{R}$, by saying that $f \in o(g)$ if, for any $\epsilon \in \mathbb{R}$, there is some $x_\epsilon \in \mathbb{R}$ such that, for any $x > x_\epsilon$, $|f(x)|/|g(x)| \leq \epsilon$

(1) Redefine big-O notation using the hyperreals, so that you don’t need to use any existential quantification (i.e., eliminate $x_0$ and $M$ from the definition).

(2) Redefine little-o notation using the hyperreals, getting rid of $\epsilon$ and $x_\epsilon$.

(3) In addition to redefining these notations for real numbers using the hyperreals, we can also say when hyperreal numbers themselves are in big-O or little-O of one another. Define this notion.

(4) Use the above to redefine the derivative without division, using little-o notation.
Definition. Suppose \( f : \mathbb{R} \to \mathbb{R} \) is a function continuous on all of \( \mathbb{R} \), and fix some constant \( a \in \mathbb{R} \). Then we define the area function \( F(t) \) to be the function giving the area of the region bounded by the graph of \( f \), the vertical lines at \( a \) and at \( t \), and the \( x \)-axis (with area under the \( x \)-axis counted as negative).

Problem 9. Prove that the derivative of \( F(x) \) exists for all \( x \), and is equal to \( f(x) \). (Hint: consider the area of the rectangle of height \( f(x + \epsilon) \) and width \( \epsilon \).)
One last time:

Definition. A function $f : \mathbb{R} \to \mathbb{R}$ is defined to be continuous at a real number $c$ if, for any $x \in {}^*\mathbb{R}$, $f^*(x) \simeq f^*(c)$ whenever $x \simeq c$.

Problem 1. Show that this definition of continuity is the same as the $\epsilon - \delta$ definition — that for any $\epsilon > 0$, there is some $\delta > 0$ such that whenever $|x - c| < \delta$, $|f(x) - f(c)| < \epsilon$.

Problem 2. Prove the Intermediate Value Theorem — given a function $f : \mathbb{R} \to \mathbb{R}$ that is continuous on the interval $[a, b]$, and given any number $c$ such that $f(a) \leq c \leq f(b)$ or $f(b) \leq c \leq f(a)$, there is some $x \in [a, b]$ such that $f(x) = c$. (Hint: if you space out $n$ points in the interval, what can you say about where you go from below $c$ to above $c$?)
Definition. If \( f \) is defined at \( x \in \mathbb{R} \), then the real number \( L \in \mathbb{R} \) is defined to be the derivative of \( f \) at \( x \) if, for every infinitesimal \( \epsilon \), \( f(x + \epsilon) \) is also defined, and
\[
\frac{f(x + \epsilon) - f(x)}{\epsilon} \simeq L
\]
We denote this number as \( f'(x) = L \).

If you’ve seen derivatives before, this should look familiar; it’s what the normal definition wishes it could be, giving the instantaneous slope of the function \( f \) at \( x \).

Problem 3. (1) If \( f \) is differentiable at \( x \), show that \( f \) must also be continuous at \( x \).

(2) Use our definition to directly compute the derivatives of \( x^n \), \( \sin(x) \), and \( e^x \) (you can assume that \( \frac{e^\epsilon - 1}{\epsilon} \simeq 1 \) for any infinitesimal \( \epsilon \); this is one way to define \( e \)).

Problem 4. Prove the Critical Point Theorem — If \( f \) has a maximum or minimum at \( x \) on some real interval \((a, b)\), and \( f \) is differentiable at \( x \), then \( f'(x) = 0 \). (Hint: compute the derivative with both a positive and negative infinitesimal.)
Problem 5. Big-O notation is used to say that one function eventually grows no faster than another; it is formally defined, for functions $f, g : \mathbb{R} \to \mathbb{R}$, by saying that $f \in O(g)$ if there are some $x_0 \in \mathbb{R}$ and $M \in \mathbb{R}$ such that, for any $x > x_0$, $|f(x)| \leq M|g(x)|$.

Little-o notation is similar, just saying that a function eventually grows much slower than another; it is formally defined, for functions $f, g : \mathbb{R} \to \mathbb{R}$, by saying that $f \in o(g)$ if, for any $\epsilon \in \mathbb{R}$, there is some $x_\epsilon \in \mathbb{R}$ such that, for any $x > x_\epsilon$, $|f(x)|/|g(x)| \leq \epsilon$

1. Redefine big-O notation using the hyperreals, so that you don’t need to use any existential quantification (i.e., eliminate $x_0$ and $M$ from the definition).

2. Redefine little-o notation using the hyperreals, getting rid of $\epsilon$ and $x_\epsilon$.

3. In addition to redefining these notations for real numbers using the hyperreals, we can also say when hyperreal numbers themselves are in big-O or little-O of one another. Define this notion.

4. Use the above to redefine the derivative without division, using little-o notation.
Definition. Suppose $f : \mathbb{R} \rightarrow \mathbb{R}$ is a function continuous on all of $\mathbb{R}$, and fix some constant $a \in \mathbb{R}$. Then we define the area function $F(t)$ to be the function giving the area of the region bounded by the graph of $f$, the vertical lines at $a$ and at $t$, and the $x$-axis (with area under the $x$-axis counted as negative).

Problem 6. Prove the Fundamental Theorem of Calculus — the derivative of $F(x)$ exists for all $x$, and is equal to $f(x)$. (Hint: consider the area of the rectangle of height $f(x + \epsilon)$ and width $\epsilon$.

Problem 7. Prove the Inverse Function Theorem — If $f$ is continuous and either strictly increasing or strictly decreasing on $(a, b)$, and $g$ is the inverse function of $f$ on that interval, then $f$ being differentiable at $x \in (a, b)$ with $f'(x) \neq 0$ means that $g$ is differentiable at $y = f(x)$, and $g'(y) = 1/f'(x)$. 