Supplemental handout - the decimal system. (Optional reading)

- We have spent the last two weeks painstakingly constructing the basic number systems of mathematics: the natural numbers, integers, rationals, and reals. Natural numbers were simply postulated to exist, and to obey five axioms; the integers then came via (formal) differences of the natural numbers; the rationals then came from (formal) quotients of the integers, and the reals then came from (formal) limits of the rationals.

- This is all very well and good, but it does seem somewhat alien to one's prior experience with these numbers. In particular, very little use was made of the *decimal system*, in which the digits $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$ are combined to represent these numbers. Indeed, except for a number of examples which were not essential to the main construction, the only decimals we really used were the numbers 0, 1, and 2, and the latter two can be rewritten as $0{++}$ and $(0{++}){++}$.

- The basic reason for this is that *the decimal system itself is not essential to mathematics*. It is very convenient for computations, and we have grown accustomed to it thanks to a thousand years of use, but in the history of mathematics it is actually a comparatively recent invention. Numbers have been around for about ten thousand years (starting from scratch marks on cave walls), but the modern Hindi-Arabic base 10 system for representing numbers only dates from the 11th century or so. Some early civilizations relied on other bases; for instance the Babylonians used a base 60 system (which still survives in our time system of hours, minutes, and seconds, and in our angular system of degrees, minutes, and seconds). And the ancient Greeks were able to do quite advanced mathematics, despite the fact that the most advanced number representation system available to them was the Roman numeral system $I, II, III, IV, \ldots$, which was horrendous for computations of even two-digit numbers. And of course modern computing relies on binary, hexadecimal, or byte-based (base 256) arithmetic instead of decimal, while analog computers such as the slide rule do not really rely on any number representation system at all. In fact, now that computers can do the menial work of number-crunching, there is very little use for decimals in modern mathematics. (Indeed, we rarely use any numbers other than one-digit numbers or one-digit fractions (as well as e, π, i) explicitly in modern mathematical work; any more complicated numbers usually get called more generic names such as n).

- Nevertheless, the subject of decimals does deserve a supplemental handout, because it is so integral to the way we use mathematics in our everyday life, and also because we do want to use such notation as $3.14159\ldots$ to refer to real numbers, as opposed to the far clunkier "$LIM_{n\to\infty} a_n$, where $a_1 = 3.1, a_2 := 3.14, a_3 := 3.141, \ldots$".

- We begin by reviewing how the decimal system works for the positive integers, and then continue on to the reals.

* * * * *

The decimal representation of natural numbers.

- In this section we will avoid the usual convention of abbreviating $a \times b$ as ab, since this would mean that decimals such as 34 might be misconstrued as 3×4.

- **Definition.** A *digit* is any one of the ten symbols $0, 1, 2, 3, \ldots, 9$. We equate these digits with natural numbers by the formulae $0 \equiv 0$, $1 \equiv 0++$, $2 \equiv 1++$, etc. all the way up to $9 \equiv 8++$. We also define the number ten by the formula ten $:= 9++$. (We cannot use the decimal notation 10 to denote ten yet, because that presumes knowledge of the decimal system and would be circular).

- **Definition.** Let $x_0, x_1, \ldots, x_n$ be a finite sequence of numbers (real, rational, integer, or natural). We define the sum $\sum_{i=0}^n x_i$, or more informally $x_0 + x_1 + \ldots + x_n$, by the recursive formulae

$$\sum_{i=0}^{0} x_i := x_0; \quad \sum_{i=0}^{n++} x_i := (\sum_{i=0}^{n} x_i) + x_{n++}.$$

- **Definition.** A *positive integer decimal* is any string $a_n a_{n-1} \ldots a_0$ of digits, where $n \geq 0$ is a natural number, and the first digit a_n is non-zero. Thus, for instance, 3049 is a positive integer decimal, but 0493 or 0 is not. We equate each positive integer decimal with a positive integer by the formula

$$a_n a_{n-1} \ldots a_0 \equiv \sum_{i=0}^{n} a_i \times \mathrm{ten}^i.$$

- Note in particular that this definition implies that

$$10 = 0 \times \mathrm{ten}^0 + 1 \times \mathrm{ten}^1 = \mathrm{ten}$$

  and thus we can write ten as the more familiar 10. Also, a single digit integer decimal is exactly equal to that digit itself, e.g. the decimal 3 by the above definition is equal to

$$3 = 3 \times \mathrm{ten}^0 = 3$$

  so there is no possibility of confusion between a single digit, and a single digit decimal. (This is a subtle distinction, and not one which is worth losing much sleep over).

- Now we show that this decimal system indeed represents the positive integers. It is clear from the definition that every positive decimal representation gives a positive integer, since the sum consists entirely of natural numbers, and the last term $a_n \mathrm{ten}^n$ is non-zero by definition.

- **Theorem 1.** Every positive integer m is equal to exactly one positive integer decimal (which is known as the *decimal representation* of m).

- To prove Theorem 1, we shall use a variant of the principle of induction, known as the *principle of strong induction*.

- **Principle of strong induction** Let $P(m)$ be a property pertaining to a positive integer m. Suppose that for each m, we have the following implication: if $P(m')$ is true for all positive integers $m' < m$ less than m, then $P(m)$ is also true. Then we can conclude that $P(m)$ is true for all positive integers m.

- In other words: if you can deduce $P(1)$ from nothing, and you can deduce $P(2)$ from $P(1)$, and you can deduce $P(3)$ from ($P(1)$ and $P(2)$), and so forth, then you can deduce $P(n)$ for every positive integer n.

- **Proof.** We use the ordinary principle of induction to prove strong induction. For any natural number n, let $Q(n)$ denote the property "$P(m)$ is true for all positive integers m which are less than or equal to n". Then $Q(0)$ is (vacuously) true, because there are no positive integers less than or equal to zero. Now suppose inductively that $Q(n)$ is true. Then $P(m')$ is true for all $m' < n++$, which implies by hypothesis that $P(n++)$ is also true. Thus $P(m')$ is in fact true for all $m' \leq n++$, which implies that $Q(n++)$ is true. Thus by induction $Q(n)$ is true for all natural numbers n, which clearly implies that $P(m)$ is true for all positive integers m. $\square$

- **Proof of Theorem 1.** We prove this by strong induction. Let $P(m)$ denote the statement "m is equal to exactly one positive integer decimal". Suppose we already know $P(m')$ is true for all positive integers $m' < m$; we now wish to prove $P(m)$.

- First observe that either $m \geq$ ten or $m \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. (This is easily proved by ordinary induction). Suppose first that $m \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Then m clearly is equal to a positive integer decimal consisting of a single digit, and there is only one single-digit decimal which is equal to m. Furthermore, no decimal consisting of two or more digits can equal m, since if $a_n \ldots a_0$ is such a decimal (with $n > 0$) we have

$$a_n \ldots a_0 = \sum_{i=0}^{n} a_i \times \text{ten}^i \geq a_n \times \text{ten}^i \geq \text{ten} > m.$$

Now suppose that $m \geq$ ten. Then by the Euclidean algorithm we can write

$$m = s \times \text{ten} + r$$

where s is a positive integer, and $r \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Since

$$s < s \times \text{ten} \leq a \times \text{ten} + r = m$$

we can use the strong induction hypothesis and conclude that $P(s)$ is true. In particular, s has a decimal representation

$$s = b_p \ldots b_0 = \sum_{i=0}^{p} b_i \times \text{ten}^i.$$

Multiplying by ten, we see that

$$s \times \text{ten} = \sum_{i=0}^{p} b_i \times \text{ten}^{i+1} = b_p \dots b_0 0,$$

and then adding r we see that

$$m = s \times \text{ten} + r = \sum_{i=0}^{p} b_i \times \text{ten}^{i+1} + r = b_p \dots b_0 r.$$

Thus m has at least one decimal representation. Now we need to show that m has at most one decimal representation. Suppose for contradiction that we have at least two different representations

$$m = a_n \dots a_0 = a'_{n'} \dots a'_0.$$

First observe by the previous computation that

$$a_n \dots a_0 = (a_n \dots a_1) \times \text{ten} + a_0$$

and

$$a'_{n'} \dots a'_0 = (a'_{n'} \dots a'_1) \times \text{ten} + a'_0$$

and so after some algebra we obtain

$$a_0 - a'_0 = (a_n \dots a_1 - a'_{n'} \dots a'_1) \times \text{ten}.$$

The right-hand side is a multiple of ten, while the left-hand side lies strictly between $-\text{ten}$ and $+\text{ten}$. Thus both sides must be equal to 0. This means that $a_0 = a'_0$ and $a_n \dots a_1 = a'_{n'} \dots a'_1$. But by the previous arguments, we know that $a_n \dots a_1$ is a smaller integer than $a_n \dots a_0$. Thus by the strong induction hypothesis, the number $a_n \dots a_0$ has only one decimal representation, which means that n' must equal n and a'_i must equal a_i for all $i = 1, \dots, n$. Thus the decimals $a_n \dots a_0$ and $a'_{n'} \dots a'_0$ are in fact identical, contradicting the assumption that they were different. $\square$

- Once one has decimal representation, one can then derive the usual laws of long addition and long multiplication to connect the decimal representation of $x + y$ or $x \times y$ to that of x or y. We won't do so here, though; you might try your hand at it if you have some energy.

- Once one has decimal representation of positive integers, one can of course represent negative integers decimally as well by using the - sign. Finally, we let 0 be a decimal as well. This gives decimal representations of all integers. Every rational is then the ratio of two decimals, e.g. $335/113$ or $-1/2$, though of course there may be more than one way to represent a rational as such a ratio, e.g. $6/4 = 3/2$.

- Since ten $=$ 10, we will now use 10 instead of ten throughout, as is customary.

* * * * *

The decimal representation of real numbers

- We need a new symbol: the *decimal point* ".".

- **Definition.** A *real decimal* is any sequence of digits, and a decimal point, arranged as

$$\pm a_n \ldots a_0 . a_{-1} a_{-2} \ldots$$

which is finite to the left of the decimal point (so n is a natural number), but infinite to the right of the decimal point, where $\pm$ is either $+$ or $-$, and $a_n \ldots a_0$ is a natural number decimal (i.e. either a positive integer decimal, or 0). This decimal is equated to the real number

$$\pm a_n \ldots a_0 . a_{-1} a_{-2} \ldots \equiv \pm 1 \times (a_n \ldots a_0 + LIM_{N \to \infty} \sum_{i=1}^{N} a_{-i} \times 10^{-i}).$$

- First, we check that this definition really does give a real number:

- **Lemma 2.** The sequence $x_1, x_2, \ldots$ defined by

$$x_N := \sum_{i=1}^{N} a_{-i} \times 10^{-i}$$

gives a Cauchy sequence of rationals. Furthermore, if $x := LIM_{N \to \infty} x_N$, then

$$x_N \leq x \leq x_N + 10^{-N}.$$

- **Proof.** Clearly each x_N is a rational number. Now let us consider the expression $|x_n - x_m|$ where $n, m \geq N$. By symmetry we may assume that $n \geq m$. Then

  $$x_n - x_m = \sum_{i=n+1}^{m} a_{-i} \times 10^{-i},$$

  (why? to rigorously prove this, one needs associativity and commutativity of addition, together with induction) and thus

  $$0 \leq x_n - x_m \leq \sum_{i=n+1}^{m} 9 \times 10^{-i}.$$

  But by the geometric series formula

  $$\sum_{i=0}^{k} ar^i = a(r^{k+1} - 1)/(r - 1)$$

  for $r \neq 1$ (this formula is easily proved by induction), we thus have

  $$0 \leq x_n - x_m \leq 10^{-n} - 10^{-m}.$$

  In particular we have

  $$|x_n - x_m| \leq 10^{-N} \text{ whenever } n, m \geq N.$$

  Since for any ε we can find an N such that $10^{-N} \leq \varepsilon$, we thus see that $(x_n)_{n=1}^{\infty}$ is a Cauchy sequence, and thus has a limit x. Also, since

  $$0 \leq x_n - x_N \leq 10^{-n} \text{ for all } n \geq N,$$

  we see after taking limits that

  $$0 \leq x - x_N \leq 10^{-N}$$

  as desired. □

- Next, we show that every real number has at least one decimal representation:

- **Theorem 3.** Every real number x has at least one decimal representation $\pm a_n \ldots a_0.a_{-1}a_{-2}\ldots$.

- **Proof**. We first note that $x = 0$ has the decimal representation $0.000\ldots$. Also, once we find a decimal representation for x, we automatically get a decimal representation for $-x$ by changing the sign $\pm$. Thus it suffices to prove the theorem for real numbers x.

- Let $n \geq 0$ be any natural number. From the Archimedean property we know that there is a natural number M such that $M \times 10^{-n} > x$. Since $0 \times 10^{-n} \leq x$, we thus see that there must exist a natural number s_n such that $s_n \times 10^{-n} \leq x$ and $s_n ++ \times 10^{-n} > x$. (If no such natural number existed, one could use induction to conclude that $s \times 10^{-n} \leq x$ for all natural numbers s, contradicting the Archimedean property).

  Now consider the sequence $s_0, s_1, s_2, \ldots$. Since we have

$$s_n \times 10^{-n} \leq x < (s_n + 1) \times 10^{-n}$$

  we thus have

$$(10 \times s_n) \times 10^{-(n++)} \leq x < (10 \times s_n + 10) \times 10^{-(n++)}.$$

  On the other hand, we have

$$s_{n++} \times 10^{-(n++)} \leq x < (s_{n++} + 1) \times 10^{-(n++)}$$

  and hence we have

$$10 \times s_n < s_{n++} + 1 \text{ and } s_{n++} \leq 10 \times s_n + 10.$$

  From these two inequalities we see that we have

$$10 \times s_n \leq s_{n++} \leq 10 \times s_n + 9$$

  and hence we can find a digit a_{n++} such that

$$s_{n++} = 10 \times s_n + a_n$$

  and hence

$$s_{n++} \times 10^{-(n++)} = s_n \times 10^{-n} + a_{n++} \times 10^{-(n++)}.$$

From this identity and induction, we can obtain the formula

$$s_n \times 10^{-n} = s_0 + \sum_{i=0}^{n} a_i \times 10^{-i}.$$

By Lemma 2, the right-hand side is a Cauchy sequence in n (clearly the addition of s_0 does not affect this property). Thus, taking formal limits of both sides, we obtain

$$LIM_{n\to\infty} s_n \times 10^{-n} = s_0 + LIM_{n\to\infty} \sum_{i=0}^{n} a_i \times 10^{-i}.$$

However, we have $s_n \times 10^{-n} \leq x$ for all n, thus

$$LIM_{n\to\infty} s_n \times 10^{-n} \leq x$$

(this almost, but not quite, follows from Corollary 22 of Week 2 notes (what is the problem?); it will however follow from material in the Week 3 notes. This will not be circular since we will not use the decimal system in the rest of the course (except as examples)). Similarly, since

$$s_n \times 10^{-n} \geq x - 10^{-n}$$

one can again take limits of both sides (assuming the material from Week 3 notes) to obtain

$$LIM_{n\to\infty} s_n \times 10^{-n} \geq x.$$

Thus we have

$$x = LIM_{n\to\infty} s_n = s_0 + LIM_{n\to\infty} \sum_{i=0}^{n} a_i \times 10^{-i}.$$

Since s_0 already has a positive integer decimal representation by Theorem 1, we thus see that x has a decimal representation. $\square$

- There is however one slight flaw with the decimal system: it is possible for one real number to have two decimal representations.

- **Proposition 4.** The number 1 has two different decimal representations: $1.000\ldots$ and $0.999\ldots$.

- **Proof.** The representation $1 = 1.000\ldots$ is clear. Now let's compute $0.999\ldots$. By definition, this is the limit of the Cauchy sequence

$$0.9, 0.99, 0.999, 0.9999, \ldots.$$

But this sequence has 1 as a formal limit, as shown in the Week 2 notes. $\square$

- It turns out that these are the only two decimal representations of 1; we leave this as an exercise for the reader. In fact, as it turns out, all real numbers have either one or two decimal representations - two if the real is a terminating decimal, and one otherwise. We won't prove this, as it is a little long and boring, but you might try it yourself if you're interested.