SEMINAR NOTES

Infinite ordinal and cardinal numbers

A set X is **countable** if there is a 1-1 function from X into the set $\mathbb{N}^+$ of positive integers, and *denumerable* if there is a 1-1 function from X *onto* $\mathbb{N}^+$. Another way to put it is that X is countable iff it can be listed in an ordinary sequence, $\langle x_1, x_2, \ldots \rangle$ with repetitions allowed, and denumerable iff it can be listed in the same way without allowing repetitions.

One of Cantor's discoveries is that $\mathbb{Q}$ is countable. It can be listed $\langle q_n : n \in \mathbb{N}^+ \rangle$, so that each $q \in \mathbb{Q}$ is encountered after finitely many steps. [To do it, list the fractions k/n in lowest terms, with positive denominator (convention: $0/1$ is in lowest terms), and with k/n coming after j/m whenever $|j| + m < |k| + n$, while if the two sums are equal then we list them in order of their numerators. Of course, for each natural number i there are only finitely many k/n with $|k| + n = i$ and n positive, so this is easily done.]

A far more important result of Cantor's is that the set of real numbers is uncountable: it is "too big" to list in a sequence in such a way that every real number is encountered after finitely many terms of the sequence. Modern set theory (and also much of modern mathematics) can be said to have started with this discovery, which was revolutionary because it showed that there is more than one infinity. In fact, there are enormously many infinite cardinal numbers, far more than the finite cardinals (natural numbers) and indeed far more than the number of elements in any set, no matter how large.

Cantor introduced both infinite cardinal numbers and infinite ordinal numbers. Both of these concepts involve well-ordering.

Definition 1. A binary relation $\preccurlyeq$ on a set X is a *well-ordering* each nonempty $A \subset X$ has a *minimum* element. This means an element $m \in A$ such that $m \preccurlyeq a$ for all $a \in A$, and such that if $a \in A$ and $a \preccurlyeq m$ then $a = m$.

The usual definitions of "well-ordering" are redundant: Definition 1 is already enough to show that $\preccurlyeq$ is reflexive (because $\{x\}$ has a minimum element), anti-symmetric (because $\{x, y\}$ has only one minimum element), transitive (compare the minumum elements of $\{x, y\}$ and $\{y, z\}$ and $\{x, z\}$) and a total order, meaning for all x, y in X either $x \preccurlyeq y$ or $y \preccurlyeq x$ (same reason as for anti-symmetry).

But such abstract algebraic treatments of well-ordering don't really convey the flavor of the concept. I will do something similar to what Descartes did when he took an ordinary plane and labeled its points by ordered pairs of real numbers, and what differential geometers do when they introduce coordinate systems on a surface, again associating ordered pairs of real numbers with the points on portions of the surface.

Every nonempty well-ordered set has a least element; label it with a 0. Throw this away; if what is left is nonempty, label its smallest element with a 1; throw this away, and if what is left is still nonempty, label its smallest element with a 2. Continue the process as long as you still have a nonempty set: if the last element you considered was labeled with k, and you got a nonempty set when you threw this away, its least element gets the label $k + 1$.

You might object that this makes perfect sense if k is a finite number, but what about infinite numbers? Well, the labels are the ordinal numbers, and the first infinite ordinal number is named (not just labeled, but *named*) ω. This is the *label* you give the minimum element of what's left after you've thrown away from the well-ordered set X all the elements that you labeled with finite numbers, assuming there is still something left of X. The next ordinal number is *named* $\omega + 1$, then the next is named $\omega + 2$, and so on. After all the $\omega + n$ ($n \in \mathbb{N}^+$) comes $\omega + \omega$, which is traditionally called $\omega \cdot 2$ or simply $\omega 2$. [I never could understand why this order was chosen rather than 2ω.]

Taking stock of what we have so far, our labels read like this:

$$0,\ 1,\ 2, \ldots, \omega,\ \ \omega+1,\ \ \omega+2, \ldots, \omega 2$$

and then come

$$\omega 2+1, \quad \omega 2+2, \quad \text{and after all the } \omega 2+n,\ n \in \mathbb{N}^+ \text{ comes } \omega 2+\omega = \omega 3, \quad \omega 3+1, \quad \ldots$$

$\ldots$ after all the $\omega \cdot 3 + n$, n a natural number, comes $\ldots$ well, you get the picture! After we've run through all the $\omega k + n$ (k, n natural numbers) we come to $\omega\omega = \omega^2$. The labels go on indefinitely, but this will do for the moment.

Note that I haven't defined ordinal numbers in general, nor will I in this note. There is a technical precise definition of "ordinal number" that has been standard since John von Neumann introduced it, but it is not especially conducive to intuition about the ordinals; even less so, than what I wrote at the beginning about well-ordered sets is conducive to intuition about *them*. But I hope that what I've said up to now helps your intuition about the following fact:

Theorem 1. *Given any two well-ordered sets* $(X, \precsim)$ *and* $(Y, \preceq)$*, either* X *is order-isomorphic to* Y*, or* X *is order-isomorphic to an initial segment of* Y*, or* Y *is order-isomorphic to an initial segment of* X*.*

[A subset A of a totally ordered set $(Z, \preccurlyeq)$ is said to be an **initial segment** of Z if whenever $a \in A$ and $z \preccurlyeq a$ then $z \in A$ also.] Since the above labeling only made use of the well-ordering, it is easy to see how the order-isomorphism **must** begin. Yes, there is no choice about it: you must match up the elements of X and Y that were labeled 0, then the ones that were labeled 1, and so on. There are enough ordinal numbers so that we never run out of labels, but if this fact sounds mysterious to you, just think of labeling the elements of X with the elements of Y until you either run out of labels (in which case Y is order-isomorphic to an initial segment of X) or you run out of elements of X (in which case X is order-isomorphic to an initial segment of Y) or both happen at the same time (in which case X and Y are order-isomorphic).

The members of $\mathbb{N}^+$ double as both ordinal and cardinal numbers. The infinite cardinals are given ordinal subscripts and conform to the labeling above. The cardinal number that expresses the number of elements in $\mathbb{N}^+$ is named $\aleph_0$ ("aleph-nought" or "aleph-zero"). If κ is a cardinal number then 2^κ denotes the number of elements in $\mathcal{P}(X)$ for some (hence all) sets X of cardinality κ. Here $\mathcal{P}(X)$ denotes the collection of all subsets of X.

Himalayan Expedition, Phase 1

One of Cantor's earliest discoveries was that the real line $\mathbb{R}$ has $2^{\aleph_0}$ elements; that is, there are just as many real numbers as there are subsets of $\mathbb{N}^+$. Cantor also showed, in general, that $\mathcal{P}(X)$ always has *more* elements than X, and so we can get a never-ending list of higher cardinals starting with $\aleph_0$:

$$\aleph_0, \quad 2^{\aleph_0}, \quad 2^{2^{\aleph_0}}, \quad 2^{2^{2^{\aleph_0}}}, \ldots$$

But how do we know we aren't skipping cardinals here? The short answer is, we don't! In fact, right at the first gap, between $\aleph_0$ and $2^{\aleph_0}$, we confront what is sometimes called Cantor's Continuum Problem. Cantor tried to prove there is no cardinal number skipped between these two; this claim is known as Cantor's **Continuum Hypothesis (CH)**. To discuss this and other matters efficiently, we need to assume something Cantor took for granted:

The Well Ordering Principle (WO) Every set can be given a well-ordering.

This principle is equivalent, in the presence of the other generally accepted axioms for set theory (commonly known as ZFC), to the following innocent-seeming axiom:

The Axiom of Choice (AC) Given any collection $\mathcal{A}$ of nonempty disjoint sets [*disjoint* means that if A_1 and A_2 are distinct members of $\mathcal{A}$, then $A_1 \cap A_2 = \emptyset$], then there is a set Z that meets every member of $\mathcal{A}$ in exactly one element.

This axiom is generally accepted nowadays as being a self-evident fact about sets. But in the early days of set theory it was controversial. Unlike with the other ZFC axioms, the AC asserts the existence of a set that is not given by a formula. This is also true of WO, and in fact, no one has come up with a formula for a well-ordering of the real line or even a well-ordering of some uncountable subset of the real line. We do know that any such well-ordering has to be different from the usual ordering; that is, every subset of $\mathbb{R}$ that is well-ordered in the usual ordering is countable. This is because, if X is a subset of $\mathbb{R}$ that is well-ordered by the usual order $\leq$, then each $x \in X$ has an immediate successor in X (namely, the minimum element of $\{x' \in X : x < x'\}$) and there are infinitely many rational numbers q such that $x < q < x'$. Distinct pairs like $\{x, x'\}$ go with disjoint sets of rationals, and, as we saw, there are only countably many rationals altogether.

Assuming WO, we can characterize the first uncountable cardinal number, which is denoted $\aleph_1$, as follows: let X be an uncountable set, and let $\preccurlyeq$ be a well-ordering of X. There are two cases:

Case 1. There is at least one $x \in X$ that is preceded by uncountably many other elements of X wrt $\preccurlyeq$. In this case, take the least such x and label it ω_1. Then $\{y \in X : y \prec \omega_1\}$ is an uncountable set that falls (with a change of notation) into:

Case 2. Every x in X has only countably many predecessors.

Take any $(X, \preccurlyeq)$ that falls into Case 2. The cardinal number $\aleph_1$ is the one that denotes the number of elements of X.

Note the analogy with $\mathbb{N}^+$, which is infinite, and yet each of its elements is preceded by only finitely many other elements.

In general, given a set Y whose cardinal number is $\aleph_\alpha$, the next cardinal number, denoted $\aleph_{\alpha+1}$, can be characterized as follows. Start with a set X whose cardinality is strictly greater than that of X; as indicated above, $\mathcal{P}(Y)$ will do nicely. Invoking WO, let $\preccurlyeq$ be a well-ordering on X. The two cases are similar to those above:

Case I. There is at least one $x \in X$ that is preceded by more than $\aleph_\alpha$-many other elements of X wrt $\preccurlyeq$. In this case, take the least such x and label it $\omega_{\alpha+1}$. Then $\{y \in X : y \prec \omega_{\alpha+1}\}$ is a set with more than $\aleph_a$-many elements, and it falls into:

Case II. Every x in X has no more than $\aleph_\alpha$ predecessors.

Take any $(X, \preccurlyeq)$ that falls into Case II. The cardinal number $\aleph_{\alpha+1}$ is the one that denotes the number of elements of X. It is easy to show that there can be no cardinal number strictly between $\aleph_\alpha$ and $\aleph_{\alpha+1}$.

When we get to ordinals not of the form $\alpha + 1$, we need a different definition. The first such nonzero ordinal is ω, and to produce a set with exactly $\aleph_\omega$ elements, we let X_n be any set of exactly $\aleph_n$ elements; then the union of all the X_n is a set of exactly $\aleph_\omega$ elements. Note that it is not necessary for the sets X_n to be disjoint: obviously, the union has more elements than any of the X_n, and it is easy to show that there cannot be any set that has more than all the X_n and less than their union. Similarly, given any set $\mathcal{K}$ of cardinal numbers, we take for each $\kappa \in \mathcal{K}$ a set X_κ of cardinality κ, and then $\sup \mathcal{K}$ is the number of elements in $\bigcup\{X_\kappa : \kappa \in \mathcal{K}\}$.

And so, the infinite cardinals are arrayed in their natural order like this:

$$\aleph_0,\ \aleph_1,\ \aleph_2, \ldots, \aleph_\omega,\ \aleph_{\omega+1}, \quad \aleph_{\omega+2}, \ldots \aleph_{\omega 2}\ \ldots \aleph_{\omega 3}\ \ldots\ \ldots \aleph_{\omega\omega} = \aleph_{\omega^2} \ldots$$

but $\aleph_\omega$ is also denoted sometimes as $\aleph_{\aleph_0}$, because most set theorists identify the ordinal ω with the cardinal $\aleph_0$. They keep both notations and use one or the other depending on what they want to emphasize. Similarly, the ordinal ω_1, which played a labeling role above, is generally identified with the cardinal number $\aleph_1$. The list above is still very far from reaching even $\aleph_{\aleph_1}$: there are only countably many cardinals on that list even if one fills in all the places where there is an ellipsis (...). But before we go higher, let me get back to Cantor's axiom CH.

In light of an early theorem of Cantor's mentioned above, CH is equivalent to the axiom $2^{\aleph_0} = \aleph_1$. We now know that this axiom is independent of the usual axioms of set theory. So is the **Generalized Continuum Hypothesis (GCH)**, which states that $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ for all infinite cardinal numbers $\aleph_\alpha$. Specifically, Gödel showed in the late 1930's that the GCH is consistent if ZFC is consistent, while Cohen showed in 1963 that CH cannot be proven by the ZFC axioms if these are consistent. Not long after Cohen, others showed that it is consistent for $2^{\aleph_0}$ to be *any of the alephs* $\aleph_\alpha$ except for those which are the supremum of a countable collection of smaller alephs. Thus $\aleph_n$ could be $2^{\aleph_0}$ for any $n \in \mathbb{N}^+$ but $\aleph_\omega$ cannot, because it is the supremum of the countably many numbers $\aleph_n$ as n runs over $\mathbb{N}^+$. In general, any infinite successor cardinal (that is, a cardinal number of the form $\aleph_{\alpha+1}$) could be assumed to equal $2^{\aleph_0}$ without fear of contradiction (unless ZFC is inconsistent, but hardly anyone seriously thinks it might be inconsistent). It is also consistent for $2^{\aleph_0}$ to equal $\aleph_{\omega_1}$, also denoted $\aleph_{\aleph_1}$.

The placement of $2^{\aleph_0}$ is relevant to an old theme relating set theory to measure theory. A **measure over a set** X is a function μ whose range is a subset of $\mathbb{R}^+ \cup \{0\} \cup \{+\infty\}$; whose domain is a collection of sets $\mathcal{A}$ known as a σ-algebra ($X \in \mathcal{A}$, $A_1 \setminus A_2 \in \mathcal{A}$ whenever A_1 and A_2 are in $\mathcal{A}$, and the union of any countable subcollection of $\mathcal{A}$ is a member of $\mathcal{A}$) and which has the following two properties: $\mu(0) = 0$, and if $\{A_n : n \in \mathbb{N}^+\}$ is a disjoint subcollection of $\mathcal{A}$ then

$$\mu(\bigcup_{n=1}^{\infty} A_n) \quad = \quad \sum_{n=1}^{\infty} \mu(A_n) \qquad (\text{ countable additivity}).$$

The measure μ is said to be *over* X and *on* $\mathcal{A}$.

A measure μ is **nontrivial** if $\mu(X) > 0$, and **a probability measure** if $\mu(X) = 1$. It is **nonatomic** if $\mu(\{x\}) = 0$ for all $x \in X$. For example, Lebesgue measure over $[0, 1]$ is a nonatomic probability measure. Many measure theorists use a different definition of "nonatomic" but as far as the following potent axiom is concerned, they are equivalent.

Axiom M. *There is a nonatomic probability measure on $\mathcal{P}(X)$ for some set X.*

"probability measure" can be replaced by a more modest concept:

Lemma. *Axiom M is equivalent to the axiom that there exists a nonatomic measure on $\mathcal{P}(X)$ such that $0 < \mu(Y) < +\infty$ for some $Y \subset X$.*

Proof. Define ν so that

$$\nu(A) = \frac{\mu(A \cap Y)}{\mu(Y)}.$$

Then ν is a probability measure on $\mathcal{P}(X)$. $\square$

In his doctoral dissertation, Stanislaw Ulam (who was probably better known as a nuclear physicist than as a mathematician) proved a striking theorem that I have taken the liberty of calling:

Ulam's Dichotomy. *Assume Axiom M. Then EITHER:*

(1) there is $Z \subset X$ such that $\mu(Z) > 0$ and for $A \subset Z$, $\mu(A) \neq 0$ implies $\mu(A) = \mu(Z)$ OR

(2) X has no more than $2^{\aleph_0}$ elements, and there is a measure on all of $\mathcal{P}(\mathbb{R})$ extending Lebesgue measure.

I have capitalized EITHER and OR because (1) and (2) are mutually exclusive for the same X, as Ulam showed.

Ulam's paper appeared in *Fundamenta Mathematicae,* volume 16 (1930). By a rare stroke of luck, it is still in our math library, whereas much more recent journals have been carted off to the Annex. The reason is that volumes 15 and 16 were bound together and volume 15 contains the index for the first fifteen volumes. The bad news is that the paper is in German. But for someone like me, whose reading knowledge of mathematical German is quite good, it's a rare treat to read the paper. It is clearly written and well motivated, and besides Ulam's Dichotomy it contains two other remarkable results.

Theorem 2. *If (2) of the dichotomy holds, then there exists an uncountable* $\aleph_\mu \leq 2^{\aleph_0}$ *such that* $\aleph_\mu$ *is not of the form* $\aleph_{\alpha+1}$*, nor it is the supremum of fewer than* $\aleph_\mu$ *cardinal numbers less than* $\aleph_\mu$*.*

Theorem 3. *If (1) of the dichotomy holds, then there exists an uncountable* $\aleph_\tau$ *less than or equal to the cardinality of* X*, such that if* $\aleph_\alpha < \aleph_\tau$ *then also* $2^{\aleph_\alpha} < \aleph_\tau$*, and* $\aleph_\tau$ *is not the supremum of fewer than* $\aleph_\tau$ *cardinal numbers less than* $\aleph_\tau$*.*

Theorem 2 has the immediate corollary that Axiom M implies CH is false, but it implies much more than that. It implies that the number of real numbers, $2^{\aleph_0}$, is enormously bigger than any of the cardinal alephs mentioned above, including $\aleph_{\aleph_1}$, even though there are uncountably many cardinal numbers just between $\aleph_{\aleph_0}$ and $\aleph_{\aleph_1}$. In fact, Theorem 2 implies that $\aleph_\mu = \aleph_{\aleph_\mu}$ — $\aleph_\mu$ is an aleph whose subscript is equal to itself. And even this barely begins to scratch the surface of how enormous $2^{\aleph_0}$ is if (2) of Ulam's Dichotomy holds. Even Theorem 2 barely begins to scratch the surface. It took over thirty years to go beyond Theorem 2, but we now know that Theorem 2 is only the base camp, so to speak, in an enormously long climb towards $2^{\aleph_0}$ if Axiom M holds.

As for (1) of Ulam's Dichotomy, it is compatible with CH, but even the smallest cardinal number that fits the description of $\aleph_\tau$ is so enormous that we are really only at the base camp for it even with Theorem 3. Phase 2 of our Himalayan Expedition will give us a much better idea of how high we have to climb before even getting a good look at this high Himalayan peak, properly referred to as "the first uncountable measurable cardinal."