6. The transfinite ordinals*

6.1. Beginnings

It seems that Cantor was led to the discovery of Set Theory by consideration of a problem in Fourier analysis, which deals with series of the following form:

$$\frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots$$

It is extraordinarily useful to be able to express a function $f : [-\pi \to \pi] \to \mathbb{R}$ in this form, and, if possible, it is nice if the expression is unique.

In 1870, Cantor had a proof of a uniqueness theorem, which we shall call Theorem 0:

Theorem 0 *Suppose that*

$$f(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots$$

$$= \frac{c_0}{2} + c_1 \cos x + d_1 \sin x + c_2 \cos 2x + d_2 \sin 2x + \cdots$$

for all values of x. Then for all n, $c_n = a_n$ and $d_n = b_n$.

This theorem was worth having in itself, but Cantor discovered he could do better; the two trigonometric series did not have to be equal for all values of x; it was enough if they were equal for all values of x not belonging to some exceptional set P, provided that P were, in some sense, small enough. To express precisely this notion of "small enough", Cantor defined what is now known as the *Cantor-Bendixson derivative* of a subset of $\mathbb{R}$ (or, more generally, any topological space):

Definition *Suppose that $P \subseteq \mathbb{R}$. Its Cantor-Bendixson derivative is the set P' of all $x \in \mathbb{R}$ such that for all open sets $U \ni x$, $U \cap P$ is infinite.*

So P' is the set of all limits of sequences of distinct points from P. (In $\mathbb{R}$, we could replace general open sets U with open intervals $(x-\epsilon, x+\epsilon)$ without changing the definition materially.)

This notion allowed Cantor to prove Theorem 1:

Theorem 1 *Suppose that P is a subset of $\mathbb{R}$ having the property that $P' = \varnothing$, and that*

$$f(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots$$

$$= \frac{c_0}{2} + c_1 \cos x + d_1 \sin x + c_2 \cos 2x + d_2 \sin 2x + \cdots$$

for all values of $x \notin P$. Then for all n, $c_n = a_n$ and $d_n = b_n$.

Indeed, the process by which he derived Theorem 1 from Theorem 0 also allowed him to derive Theorem 2 from Theorem 1:

---

* The material in this handout is, in part, a simplification of material from J. Dauben, *Georg Cantor*; fuller details are available in that book.

<u>Theorem 2</u> *Suppose that P is a subset of $\mathbb{R}$ having the property that $P'' = \varnothing$, and that*

$$f(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots$$

$$= \frac{c_0}{2} + c_1 \cos x + d_1 \sin x + c_2 \cos 2x + d_2 \sin 2x + \cdots$$

for all values of $x \notin P$. Then for all n, $c_n = a_n$ and $d_n = b_n$.

and Theorem 3 from Theorem 2:

<u>Theorem 3</u> *Suppose that P is a subset of $\mathbb{R}$ having the property that $P''' = \varnothing$, and that*

$$f(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots$$

$$= \frac{c_0}{2} + c_1 \cos x + d_1 \sin x + c_2 \cos 2x + d_2 \sin 2x + \cdots$$

for all values of $x \notin P$. Then for all n, $c_n = a_n$ and $d_n = b_n$.

and, more generally, Theorem $n + 1$ from Theorem n:

<u>Theorem n</u> *Suppose that P is a subset of $\mathbb{R}$ having the property that $P^{(n)} = \varnothing$, and that*

$$f(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots$$

$$= \frac{c_0}{2} + c_1 \cos x + d_1 \sin x + c_2 \cos 2x + d_2 \sin 2x + \cdots$$

for all values of $x \notin P$. Then for all n, $c_n = a_n$ and $d_n = b_n$.

where we define $P^{(n)}$ by recursion on n so that $P^{(0)} = P$ and $P^{(n+1)} = (P^{(n)})'$.

Thus we now have a sequence of theorems, each of which is better than the previous one, because the condition on P is weaker; so the set of exceptional points, on which the trigonometric series can "go wrong", is allowed to be larger. (The reason that the conditions are getting weaker is that for any P, $P' \supseteq P'' \supseteq P''' \supseteq \cdots$. If you have studied some topology, see if you can figure out why.)

But Cantor wasn't finished there. Can we write down a condition weaker than all the conditions $P' = \varnothing$, $P'' = \varnothing$, $P''' = \varnothing$, ..., $P^{(n)} = \varnothing$? Yes we can: what about

$$\bigcap_{n \in \mathbb{N}} P^{(n)} = \varnothing?$$

Using this even weaker condition, Cantor was able to prove yet another theorem which, using notation which it took Cantor some years to settle on, we shall call Theorem ω:

<u>Theorem ω</u> *Suppose that P is a subset of $\mathbb{R}$ having the property that $P^{(\omega)} = \varnothing$, and that*

$$f(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots$$

$$= \frac{c_0}{2} + c_1 \cos x + d_1 \sin x + c_2 \cos 2x + d_2 \sin 2x + \cdots$$

for all values of $x \notin P$. Then for all n, $c_n = a_n$ and $d_n = b_n$.

where $P^{(\omega)} = \bigcap_{n \in \mathbb{N}} P^{(n)}$.

Now if P' is the first derivative of P, P'' the second derivative, and so on, it is surely natural to refer to $P^{(\omega)}$ as the ωth derivative. In grammar, the words *first, second, third* and so on are called ordinal numbers; so surely the ω in the notation $P^{(\omega)}$ is playing the same role; it is an infinite ordinal number. Similarly, when we refer to Theorem 3, we do not think that there are three of this theorem; we are saying that it is the *third* in some sequence: so in "Theorem 3", the label *3* is really an ordinal; so we can regard the ω in "Theorem ω" in the same way.

Why stop there? How about the theorem concerning $(P^{(\omega)})'$? This is derived from Theorem ω in the same way that we derive Theorem 3 from Theorem 2 or Theorem 57 from Theorem 56, so it's tempting to call it Theorem $\omega + 1$:

Theorem $\omega + 1$ *Suppose that P is a subset of $\mathbb{R}$ having the property that $P^{(\omega+1)} = \varnothing$, and that*

$$f(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots$$

$$= \frac{c_0}{2} + c_1 \cos x + d_1 \sin x + c_2 \cos 2x + d_2 \sin 2x + \cdots$$

for all values of $x \notin P$. Then for all n, $c_n = a_n$ and $d_n = b_n$.

where $P^{(\omega+1)} = (P^{(\omega)})'$.

In the same way, we define $P^{(\omega+2)}$, $P^{(\omega+3)}$ and so on, and prove Theorem $\omega + 2$, Theorem $\omega + 3$ and so on.

And then there's a theorem which is better than all of these. If we define

$$P^{(\omega+\omega)} = \bigcap_{n \in \mathbb{N}} P^{(\omega+n)},$$

then we can prove a Theorem $\omega + \omega$; and after that a Theorem $\omega + \omega + 1$, $\omega + \omega + 2$, ..., and then eventually Theorem $\omega + \omega + \omega$, $\omega + \omega + \omega + 1$, $\omega + \omega + \omega + 2$, ..., $\omega + \omega + \omega + \omega$, ..., $\omega + \omega + \omega + \omega + \omega$, ...

There is clearly going to be a theorem better than all of these; but we need some new notation to give it a name. If we write $\omega + \omega$ as $\omega.2$ and $\omega + \omega + \omega$ as $\omega.3$ and so on, the next theorem in the list is perhaps going to be $\omega.\omega$, which we might write as ω^2. That is going to be followed immediately by $\omega^2 + 1$, $\omega^2 + 2$, ..., $\omega^2 + \omega$, $\omega^2 + \omega + 1$, ..., $\omega^2 + \omega.2$, ..., $\omega^2 + \omega.3$, ..., $\omega^2 + \omega^2 = \omega^2.2$, ...

6.2. Classifying the ordinals

It is reasonable to classify the natural numbers into two classes:

1. 0,
2. Everything else.

All natural numbers other than zero are the successor of something; that is, they are of the form $n + 1$ for some n. So, we can rewrite the above list; every natural number belongs to one of the two classes

1. 0,
2. successors $n + 1$.

Looking at the ordinals (finite and infinite), we see that these two cases do not exhaust the possibilities. What is ω? It obviously isn't zero. It also isn't of the form $\alpha + 1$, because if $\alpha < \omega$, then α is a natural number, and then so is $\alpha + 1$. So ω belongs to a third class.

So we classify the ordinals as follows:

1. 0,
2. successors $\alpha + 1$,
3. everything else.

The term we use for members of the third class is *limit ordinals*.

6.3. Ordinal arithmetic

We've been deriving our notation for ordinal numbers by using arithmetical operations; evidently it would be as well to have some idea of how these work.

The principles of ordinal arithmetic are fundamentally different from those of cardinal arithmetic, in that they are all about arranging things in an order. Let's start with finite ordinals. What is addition, in this context?

Let's consider the English king Henry VIII. Did we obtain him by taking three Henrys, then five more, and adding them together? Not really: Henry VIII was just one man after all. But what we can say is that he was the *fifth* king called Henry after Henry *III*. So in ordinal arithmetic the equation

$$3 + 5 = 8$$

means: advance three steps along your list; then advance five more; that's the same as going straight to the eighth position. And notice again that this is all about order. And notice also that it's not quite instantly obvious that this operation is commutative.

Now let's do some addition with infinite ordinals. Let's calculate $\omega + 2$ and $2 + \omega$, and compare them. To calculate $\omega + 2$, we first count off all the ordinal numbers below ω:

$$0, \; 1, \; 2, \; 3, \; \ldots$$

and then count off an additional copy of the ordinals below 2:

$$0, \; 1, \; 2, \; 3, \; \ldots; \; 0, \; 1$$

and then count off the whole list, labelling them with the ordinal numbers in order:

$$0, 1, 2, 3, \ldots; 0, \quad 1$$
$$0, 1, 2, 3, \ldots; \omega, \omega + 1$$

and what we've created is a copy of all the ordinals below $\omega + 2$, as we would have expected.

Now let's do the sum the other way round. First count off two:

$$0, \; 1$$

then an additional ω:

$$0, \; 1; \; 0, \; 1, \; 2, \; 3, \; \ldots$$

4

and then assign ordinal labels to the entire list, in order:

$$0, 1; 0, 1, 2, 3, \ldots$$
$$0, 1, 2, 3, 4, 5, \ldots$$

This time, we've only used up the ordinals before ω. So rather unexpectedly, we find that $2 + \omega = \omega \neq \omega + 2$.

Well, perhaps it isn't so surprising. After all, ordinals, by their nature, are about order. So when we add two ordinals, we might very well expect it to matter which order we do it in.

The account of ordinal addition given here—to obtain $\alpha + \beta$, first count off a copy of (the ordinals below) α, then an additional copy of (the ordinals below) β, and then see what the result looks like—is the account best suited to forming intuitions about ordinal addition. However, the definition given in lectures is better for proving theorems about it. To really understand the ordinals, it helps to have both.

How about ordinal multiplication? In ordinal arithmetic, the equation

$$2.4 = 8$$

means "after you've counted off two for the fourth time, you'll be at the eighth position". In the context of kings of England, we count off an ordered pair of Henrys four times:

Henry I	Henry II	Henry III	Henry IV	Henry V	Henry VI	Henry VII	Henry VIII
first	second	first	second	first	second	first	second
	1st pair		2nd pair		3rd pair		4th pair

and arrive at our old friend Henry VIII.

Now let's try this with some infinite ordinals. Again, let's calculate $2.\omega$ and $\omega.2$. To calculate $\alpha.\beta$, we write out copies of α, one after the other, one for each ordinal less than β.

So to calculate $\omega.2$, first write out (the ordinals less than) 2:

$$0, 1$$

then to each assign a copy of (the ordinals less than) ω:

$$0, \qquad\qquad 1$$
$$0, 1, 2, 3, \ldots; 0, 1, 2, 3, \ldots$$

then assign ordinal labels to the result, in order:

$$0, \qquad\qquad\qquad\qquad 1$$
$$0, 1, 2, 3, \ldots; 0, \quad 1, \quad 2, \quad 3, \quad \ldots$$
$$0, 1, 2, 3, \ldots; \omega, \omega+1, \omega+2, \omega+3, \ldots$$

and we find that we've used up the ordinals less than $\omega + \omega$, as expected.

Now we repeat for $2.\omega$. This time we begin with a copy of (the ordinals less than) ω:

$$0, \ 1, \ 2, \ 3, \ \dots$$

then to each one assign a copy of (the ordinals less than) 2:

$$\begin{array}{cccc} 0 & 1 & 2 & 3 \dots \\ 0, 1; & 0, 1; & 0, 1; & 0, 1; \dots \end{array}$$

then assign ordinal labels to the result:

$$\begin{array}{cccc} 0 & 1 & 2 & 3 \dots \\ 0, 1; & 0, 1; & 0, 1; & 0, 1; \dots \\ 0, 1, & 2, 3, & 4, 5, & 6, 7, \dots \end{array}$$

This time we've used up just the natural numbers, and we find that $2.\omega = \omega \neq \omega.2$.

Again, this method of defining ordinal multiplication is better for forming intuition, but the recursive method used in lectures is better for proving theorems.

There is a similar picture for ordinal exponentiation, but it is not very helpful.

It might be a useful exercise to try out these intuitive ideas, to convince yourself why the following are true:

1. If β is a limit, then so is $\alpha + \beta$,
2. if β is a successor, then so is $\alpha + \beta$,
3. if α or β is a limit, and α and β are not zero, then $\alpha.\beta$ is a limit,
4. if α and β are both successors, then $\alpha.\beta$ is a successor,
5. ordinal addition and multiplication are associative,
6. ordinal addition and multiplication allow cancellation on the left: that is, if $\alpha + \beta = \alpha + \gamma$, then $\beta = \gamma$, and if $\alpha.\beta = \alpha.\gamma$ and $\alpha \neq 0$, then $\beta = \gamma$,
7. ordinal addition and multiplication do not allow cancellation on the right,
8. the right distributive law $\alpha(\beta + \gamma) = \alpha.\beta + \alpha.\gamma$ is true, but the left distributive law $(\alpha + \beta).\gamma = \alpha.\gamma + \beta.\gamma$ is not.

All of these can be more conveniently *proved* using the formal definitions given in lectures.

6.4. Ordinals and cardinals part company

The *finite* cardinals and ordinals are isomorphic, and so, mathematically, we tend not to bother to distinguish them. However, when we proceed to infinite cardinals and ordinals, differences emerge.

For a start, it should be obvious from the above discussion that ordinal arithmetic, and cardinal arithmetic, have totally different properties.

But there is a more important and more basic difference between the two systems.

The finite cardinals and ordinals are in one-to-one correspondence. Everyone who has tried to learn a language is familiar with this, from tables like the following:

Cardinals			Ordinals	
one	unus		first	primus
two	duo		second	secundus
three	tres		third	tertius
four	quattuor		fourth	quartus
five	quinque		fifth	quintus
six	sex		sixth	sextus
seven	septem		seventh	septimus
eight	octo		eighth	octavus
nine	novem		ninth	nonus
ten	decem		tenth	decimus

and you certainly expect the rows of the two tables to line up.

But this fails dramatically when we come to the infinite cardinals and ordinals.

To see why, let's try to see what cardinals correspond to the ordinals we've defined so far.

We begin with ω. The cardinal corresponding to it is presumably the number of ordinals less than ω, that is, the number of elements in the following sequence:

$$0,\ 1,\ 2,\ 3,\ \ldots$$

which is $\aleph_0$; that is, the list is countably infinite.

How many ordinals are there below $\omega + 1$? Well, all of the natural numbers, and one more, namely ω itself. So the corresponding cardinal is $\aleph_0 + 1$. However, in view of the equation

$$\aleph_0 + 1 = \aleph_0,$$

the cardinal number associated with $\omega + 1$ is just $\aleph_0$ again; and, applying the equation repeatedly, the same is true of $\omega + 2$, $\omega + 3$ and so on. What about $\omega.2 = \omega + \omega$? In this case, we use the equation

$$\aleph_0 + \aleph_0 = \aleph_0$$

to get the same answer. In fact, by using in addition the equation

$$\aleph_0.\aleph_0 = \aleph_0,$$

we can see that all of the infinite ordinals we have mentioned so far correspond to the same cardinal number.

So if we were to extend the tables we had earlier, we would obtain

Cardinals

one	unus
two	duo
three	tres
four	quattuor
five	quinque
six	sex
seven	septem
eight	octo
nine	novem
ten	decem
$\vdots$	$\vdots$
$\aleph_0$	

Ordinals

first	primus
second	secundus
third	tertius
fourth	quartus
fifth	quintus
sixth	sextus
seventh	septimus
eighth	octavus
ninth	nonus
tenth	decimus
$\vdots$	$\vdots$
ω	
$\omega + 1$	
$\omega + 2$	
$\vdots$	
$\omega.2$	
$\omega.2 + 1$	
$\vdots$	
$\omega.3$	
$\vdots$	

(I am not familiar with the Latin words for the infinite ordinals and cardinals.)

None of the operations of ordinal arithmetic suffices to get us an uncountable ordinal; and indeed, the first one is named using new notation as ω_1. (And the next ordinal after it is $\omega_1 + 1$, then comes $\omega_1 + 2$, $\omega_1 + 3$, …, $\omega_1 + \omega$, …)

6.5. Identifying the ordinals as sets

It is now customary to identify the ordinal α with the set of all preceding ordinals $\{\beta : \beta < \alpha\}$, and we use this standard identification in this lecture course. However, if you tried to convince your greengrocer that the natural number 0 was actually the empty set, 1 was actually $\{0\}$, and so on, then they would think you were a bit strange, and they would be right. In the same way, although the standard way of identifying the ordinals as sets (which results, in particular, in ω becoming the set of natural numbers) works and is convenient, it does not seem to me to be any more significant than that. I met the infinite cardinals and ordinals many years before I first took a set theory course, and to me it seems that the preceding intuitive discussion comes closer to what the ordinals "really are".