A Tutorial on Countable Ordinals

Thomas Forster

November 22, 2010

Contents

ABSTRACT

Cantor and derived sets. Ordinals as a recursive datatype and Ordinals as isomorphism types. Definition of *countable*, *of cofinality*. Clubsets, normal functions. Families of binary functions; notations and normal forms. Dominance; fast-growing functions $\mathbb{N} \to \mathbb{N}$. Fundamental sequences. Connections with primitive recursion, double recursion etc. Schmidt-coherence.

Prerequisites: wellfounded relations. (Also a bit of recursion theory if we can to cover everything here)

At some point must discuss what goes wrong if you try to use the Cantor Normal Form division algorithm with a base other than ω. I've never thought about it, but it goes wrong, and that is why it is never done!

Acknowledgements

I can be quite sure that much of the material below was explained to me by patient friends and colleagues—to whom I undoubtedly owe a huge debt of gratitude. Sadly, the manner in which I internalised this material was such as to render its provenance unascertainable on subsequent regurgitation, so I cannot now be entirely sure what I learned from whom! (There is even the possibility—admittedly remote—that I actually managed to work some of this out for myself!) One thing I do know is that I have profited greatly from the patience and understanding of Adrian Mathias, Harold Simmons, Martin Hyland and Stanley Wainer at least, and it is a pleasure to be able to record my endebtedness to them—and my thanks—here. It is a pleasure also to be able to record my thanks to those of my students who, trawling though these notes in the expectation of gaining enlightenment thereby, discovered instead a rat's nest of errors which they were then kind enough to give me the opportunity of silently correcting.

Notation, Background, etc

Lowercase Greek letters are used to range over ordinals. Its use in λ-calculus notwithstanding, the letter 'λ' is always liable to be a variable ranging over *limit* ordinals in the way that in A-level analysis 'x' and 'y' are ordinate and abcissa, or input and output variables. With this in mind I shall refrain from using lambda notation, using the '$\mapsto$' notation instead.

I am going to assume that you all know what a wellfounded relation is, what wellfounded induction is, and know what a recursive datatype is. The colex ordering of sequences from a set is the ordering by last difference. explain *colex*

I am also going to assume that you know a bit of first-year analysis: the rationals are countable and dense in the reals (which are not countable); there is a real between any two rationals and a rational between any two reals. The set of naturals is of size $\aleph_0$; the continuum is of size $2^{\aleph_0}$. The **Continuum Hypothesis** is the proposition that $2^{\aleph_0} = \aleph_1$. Perhaps you do not yet know what $\aleph_1$ is but this will be explained to you on page 9.

I am going to assume that you know a bit of recursive function theory, though not very much, and only in the last few pages.

0.1 Finite Objects

I'm going to assume that you have a concept of *finite object*. A set X of things is a set of finite objects iff there is a system of notation for members of X such that every member of X has a finite description. (Natural numbers are finite objects; rationals and algebraics are finite objects; reals famously are not. They are *infinite precision* objects.) The observant reader will complain that—according to this definition—any object that belongs to a countable set X can be made to be a finite object: all that one has to do is fix in advance a bijection between X and $\mathbb{N}$, and then one can point to an object by saying that it is the nth member of X according to the given enumeration. Of course life is not that simple. One does not want X to be just any old random assemblage of things, one wants it to be a set in the rather stricter sense in which one speaks of a *set of spoons*, or a *set of plates*, or a *set of rules*, or a *chess set*: X must be a family of homologous objects admitting a uniform description (or a union of finitely many such families). Further, the enumeration of this non-random collection must be in some informal sense computable. Indeed there is a useful and practical converse to this, which I impress on all my first-years. If you are presented with a natural set (not a mere assemblage) and you want to know whether or not it is countable: ask yourself: are its members finite objects? Do I have a uniform finitary system of notation for its members? If I do, it's countable—and if it doesn't it isn't. This simple heuristic is a remarkably efficacious way deciding whether or not a candidate set is countable.

The concept of finite object is not a mathematically rigorous one, but it is very important nevertheless. I have a hunch that the most sympathetic (and quite possibly the most correct) way to understand the Hilbert programme is as an endeavour to represent as much as possible of mathematics as the study of finite objects. Finitism started off as a sensible idea: ideologies always do—however crazy they turn out to be later. Look at how much progress in Mathematics involves reducing problems to finite calculations. Once you have any intuition of a difference between finite objects and infinite objects you notice that finite objects are tractable and infinite objects aren't, and progress in the study of particular kinds of mathematical objects happens when you find ways of thinking of them as finite objects. (Algebraic topology etc.; Euler's polyhedron formula is a nice example of distillation of finite information from infinite sources. Knots.) Proofs are finite objects; all of syntax is peopled with finite objects. It is not at all barmy to think that mathematics is really the study of finite objects, and that a preoccupation with trying to express everything in terms of structure of finite character is the way to go. It may be *mistaken*, but it certainly isn't *barmy*. It's mistaken beco's the aim of Mathematics is to generalise, but it's not crazy.

A set-of-finite-objects is a set equipped with enough structure for there to be a system of notation that allocates everything in the suite a description containing only finitely many symbols. The minimal conditions for this to happen seem to be for the set to be a recursive datatype of finite character, or—to put it another way— (using an encoding scheme) an *r.e.* or *semidecidable* set of naturals. This is why the (recursive) axiomatisablity of First Order Logic is so important: valid sentences of First Order Logic come equipped with proofs that are finite objects, but valid sentences of higher-order logic do not.

Given their importance, clarifying the concept of finite object is probably a good project. One way into it is to think about countable ordinals. We will see that the collection of countable ordinals is itself uncountable, and so its members cannot be thought of as finite objects. However all its proper initial segments are countable, so the inhabitants of any proper initial segment can be thought of as finite objects. But not uniformly! It was the thought that this nonuniformity could be an opening into the concept of finite object that was one of the attractions for me of the project of understanding countable ordinals.[1]

1 Ordinals as a Recursive Datatype

1.1 Cantor's discovery of ordinals

Ordinals were invented by Cantor to solve a problem in the theory of Fourier series. Although it's an interesting story I shall consider only those bits of it that are directly relevant.

A Fourier series whose every coefficient is zero is obviously the identically zero function. What about the converse? Cantor's first theorem said that if S is a Fourier series which converges to 0 everywhere then all coefficients are zero.

Obvious question: can we weaken the hypothesis by weakening 'everywhere' to 'except on a something-or-other set'. The answer is: yes, indeed we can. (Think about the Fourier series for a square wave.) Quite how far one can weaken it is a question that doesn't have a nice answer. However Cantor was able to show that "something-or-other" can be closed-countable, and he did this by transfinite induction on the rank of closed sets. It turns out that the assumption of closedness is unneccessary, as was shown by an Englishman by the name of 'Young' by a completely different method.[2] But the trip up the blind alley at least gave us ordinals.

Cantor was interested in applying to an arbitrary closed set X of reals the operation that returns its **derived set**: the set of all limit points of X. If X is closed its derived set is a subset of it. How often can one apply this operation to a closed set before one reaches either an empty set or a perfect closed set (which is a fixed point, being equal to its derived set)? The interesting point here is that since this operation is monotone decreasing with respect to $\subseteq$ it makes sense to think of transfinite iteration: one can take intersections at limit stages and carry on deriving. So the answer to the question "How often?" might not be a natural number. What sort of number is it? The answer is that it will be an **ordinal**. Ordinals are the kind of number that measures the length of precisely this sort of process: transfinite and discrete.

[1] There is an apparent paradox here (which we shouldn't really discuss at this point, for fear of frightening the horses). We shall see later that, for any countable ordinal α, every countable ordinal $\beta > \alpha$ gives us a way of thinking of α (indeed of every ordinal $< \beta$) as a finite object. But there are uncountably many countable ordinals $\beta > \alpha$ so this means that there are *uncountably* many finitary systems of notation for countable ordinals. But a finitary system of notation is itself a finite object, being a finite set of rules over a countable alphabet, so there are only countably many of them. This will be resolved later in these notes (see p 11)by the concept of a *recursive* ordinal.

[2] I stumbled upon the article in which Young proved this. It is in the same volume of the same journal as the article [12] by Hardy below!

The idea that ordinals count the length of discrete transfinite processes should be taken seriously and can be taken further. There is an addition operation on processes, written '+' with overloading, but no operation of multiplication of processes *by processes*. However there is a notion of multiplication of a process by an ordinal ("Do this α times"): a process multiplied on the right by an ordinal is another process.

Thus, if we let $\mathfrak{p}$ and $\mathfrak{s}$ be processes, and let α and β be ordinals then we have the following easy equations:

1. $\mathfrak{p} \cdot (\alpha + \beta) = \mathfrak{p} \cdot \alpha + \mathfrak{p} \cdot \beta$
2. $\mathfrak{s} \cdot (\alpha \cdot \beta) = (\mathfrak{s} \cdot \alpha) \cdot \beta$

and others like it. The effect is that processes form a module over the ordinals. In fact this could be an operational way of characterising the ordinals: as that-kind-of-number-such-that-processes-form-a-module-over-them[3]

I'm not sure how seriously this idea should be taken: granted, there seems to be a good notion of *length* of a process, but it's pretty clear that there is no good notion of *inner product* nor of *dot product* of two processes. However I have recently (re)discovered a typescript of Girard and Norman which says that things called *dilators* (which we may see later) behave like linear operators in vector spaces ... so presumably they had the same idea.

Even if all we know about ordinals is that they are the kind of number that enumerates the stages in processes like that of Cantor's we considered, we nevertheless know quite a lot about them. At any stage there is always in principle the possibility of a next stage, so the successor of an ordinal is an ordinal. But because the operation of taking-the-derived-set is monotone, there is a concept of a limit stage, so it must be that a supremum of a set of ordinals is an ordinal.

At this point I should really give a presentation of the ordinals as a recursive datatype. Unfortunately I am not in a position to do this, since locating the exact answer turns out to be a fiddlier task than I had hoped, and it may well be that there is more than one way of doing it. I shall restrict myself to making some basic but (i hope) helpful observations.

1. The idea is that the Ordinals are like the naturals with an extra constructor: `sup` of, which is applied to sets of ordinals. This makes it a higher-order recursive datatype.

2. Since distinct sets of ordinals can have the same `sup` the `sup` constructor is not free, and this is the chief source of the trouble.

3. The reader should rehearse the way in which the declaration of $\mathbb{N}$ as a rectype gives rise to the engendering relation $<_{\mathbb{N}}$ and a proof that that engendering relation is a total order, and wellfounded. Make sure you understand that. Once you do, you will be able to see what a declaration of the ordinals should look like.

[3]Pedants will delight in pointing out that modules are formed over *rings* and that the ordinals do not form a ring.

For the moment I am going to assume that all that has been dealt with.

1.2 Operations on ordinals

Now we can give some recursive definitions of the obvious operations, starting with addition.

Definition 1
$\alpha + 0 := \alpha$;
$\alpha + \mathtt{succ}(\beta) := \mathtt{succ}(\alpha + \beta)$;
$\alpha + \mathtt{sup}(X) := \mathtt{sup}\{\alpha + \beta : \beta \in X\}$.

Theorem 2 *For all ordinals α and β, $\alpha \leq_{On} \beta$ iff $(\exists\gamma)(\alpha + \gamma = \beta)$.*

The best way to prove this is use of the picture of ordinals as isomorphism classes. See section 1.3 below.

Next a little lemma we shall need later.

Lemma 3

1. $(\forall\alpha)(\forall\beta)(\alpha \leq \alpha + \beta)$
2. $(\forall\alpha)(\forall\beta)(\beta \leq \alpha + \beta)$

Proof: The proof of this falls into two cases, and the cases are different because addition on the left is different from addition on the right

Case 1: $\alpha \leq \alpha + \beta$

For each α we prove by induction on β that $\alpha \leq \alpha + \beta$.

Case 2: $\beta \leq \alpha + \beta$

For $\beta \leq \alpha + \beta$ we prove by induction on β that $(\forall\alpha)(\beta \leq \alpha + \beta)$.

Clearly $(\forall\alpha)(0 \leq \alpha + 0)$

For the successor case, assume $(\forall\alpha)(\beta \leq \alpha + \beta)$. We want $(\forall\alpha)(\beta + 1 \leq \alpha + \beta + 1)$. But clearly $\beta \leq \alpha + \beta$ iff $\beta + 1 \leq \alpha + \beta + 1$.

For the limit case let $\lambda = \mathtt{sup}X$. Let α be arbitrary. We want $\lambda \leq \alpha + \lambda$.

$\alpha + \lambda = \alpha + \mathtt{sup}X$

$\alpha + \mathtt{sup}X = \mathtt{sup}\{\alpha + \beta : \beta \in X\}$

But now (by induction hypothesis) everything β in X is $\leq$ something (to wit: $\alpha+\beta$) in $\{\alpha+\beta : \beta \in X\}$ so $\mathtt{sup}X$—which is λ—is $\leq \mathtt{sup}\{\alpha+\beta : \beta \in X\}$—which is $\alpha + \lambda$.

■

Notice that for Case 1 we did a Δ_0 induction and for Case 2 we had to do a Π_1-induction. Addition on the right is easier to reason about than addition on the left!

Lemma 4 *The ordinals $\leq \alpha$ are totally ordered by $\leq$.*

Proof: We do this by induction on α. The base case is immediate; the `succ` case is just like the inductive proof that $<_{\mathbb{N}}$ is a total order. For the limit case we exploit $(\forall \alpha)(\forall S \subset On)((\alpha < \mathtt{sup}(s)) \rightarrow (\exists \beta \in S)(\alpha \leq \beta))$. details here

If $\alpha_1 \leq \mathtt{sup}(S)$ and $\alpha_2 \leq \mathtt{sup}(S)$ then there is $\beta \in S$ with $\alpha_1 \leq \beta$ and $\alpha_2 \leq \beta$. (Indeed, by lemma **??**, β can be taken to be $\alpha_1 + \alpha_2$ or $\alpha_2 + \alpha_1$.) But then α_1 and α_2 are comparable by induction hypothesis.

■

Corollary 5 *$<_{On}$ is a wellorder.*

Proof: It's wellfounded because it is the engendering relation of a rectype. To show it's a total order consider two arbitrary ordinals α and β. By lemma 3, α and β are both $\leq \alpha + \beta$. Then by lemma 4 the ordinals below $\alpha + \beta$ are totally ordered. ■

This proof of corollary 5 is mine, though it may well have been anticipated. If so, I hope my readers will tell me. There is a proof concealed in the papers of Bourbaki [2] and Witt [21] (See Appendix 1).

> *It now seems to me that one can give a much shorter proof that $<$ is a total order. We know it is wellfounded. Consider a minimal member α_1 of $X = \{\alpha : (\exists \beta)(\alpha \neq \beta \not> \alpha \not> \beta)\}$, and then a minimal member α_2 of $\{\alpha : \alpha \neq \alpha_1 \not> \alpha \not> \alpha_1\}$. Thus α_1 and α_2 are incomparable minimal elements of X. The ordinals below α_1 form a chain A_1 and the ordinals below α_2 form a chain A_2. Now these must be the same chain, so we call it A. If A has a top element—α, say—then α_1 and α_2 must both be $\mathtt{succ}(\alpha)$. If not, they must both be $\mathtt{sup}(A)$. Either way, they are the same.*

Now we can procede to define multiplication:

Definition 6

$\alpha \cdot 0 := 0$;
$\alpha \cdot \mathtt{succ}(\beta) := (\alpha \cdot \beta) + \alpha$;
$\alpha \cdot \mathtt{sup}(X) := \mathtt{sup}(\{\alpha \cdot \beta : \beta \in X\})$;

and exponentiation:

Definition 7

$\alpha^0 := \mathtt{succ}(0)$;
$\alpha^{(\mathtt{succ}(\beta))} := (\alpha^\beta) \cdot \alpha$;
$\alpha^{(\mathtt{sup}(X))} := \mathtt{sup}(\{\alpha^\beta : \beta \in X\})$.

Given these definitions, it is clear that addition on the right, multiplication on the right and exponentiation on the right, namely, the functions $\alpha \mapsto (\beta + \alpha)$, $\alpha \mapsto (\beta \cdot \alpha)$ and $\alpha \mapsto (\beta^\alpha)$ are—for each ordinal β—*continuous* in the sense in which the ordinals are (very nearly) a chain-complete poset.

Exercise 1

1. *Give examples to show that addition and multiplication on the left are not commutative.*
2. *Give an example to show that* $\alpha \mapsto \alpha^2$ *is not continuous.*
3. *Which of the following are true for all* α, β *and* γ*?*
   (a) $(\alpha \cdot \beta)^\gamma = \alpha^\gamma \cdot \beta^\gamma$;
   (b) $\gamma^{(\alpha+\beta)} = \gamma^\alpha \cdot \gamma^\beta$;
   (c) $(\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma$;
   (d) $\gamma \cdot (\alpha + \beta) = \gamma \cdot \alpha + \gamma \cdot \beta$.

   Prove the true assertions and give counterexamples to the false assertions.

Definition 8 *If* $\langle X, R\rangle$ *and* $\langle Y, S\rangle$ *are two wellfounded binary structures then* $f : X \to Y$ *is* **parsimonious** *if, for all* $x \in X$, $f(x)$ *is an* S*-minimal* y *in* Y *such that* $(\forall x' R x)(f(x') S y)$.

Remark 9 *For every wellfounded binary structure* $\langle X, R\rangle$ *there is a unique parsimonious map to the ordinals. That is to say that* On *is a terminal object in the category of wellfounded binary structures and parsimonious maps.*

1.3 Ordinals as Order Types

If we think of On as an abstract datatype in this way it is natural to associate to each ordinal a picture consisting of a string of dots and lines, in fact: a Hasse diagram. Now Hasse diagrams are usually pictures of *posets*. This reminds us that ordinals can also be thought of as relational types (isomorphism classes) of posets: rather special posets in fact, to wit *wellorderings*.

A Wellordering is a wellfounded strict partial order. An ordinal is an isomorphism type of wellorderings. (I don't want to say it's an isomorphism class, because I really don't want to get embroiled in issues of set existence!) Thus we say that an ordinal is the *length* of a wellordering whose *type* it is. People sometimes write 'order type of' instead of 'length of'. A countable ordinal is the ordinal of a wellordering whose carrier set is countable, without loss of generality the length of a wellordering of $\mathbb{N}$. **Countable ordinal**

The following fact about ordinals is of fundamental importance. For any ordinal α, α is the order type of the set $\{\beta : \beta < \alpha\}$ of ordinals below α in the obvious ordering (of definition 35.) This fact is so cute that it has become the basis of the standard implementation of ordinal arithmetic into set theory. In this implementation (due to Von Neumann) each ordinal is simply taken to *be* the set of ordinals below it. more details here

There is a slight niggle over this, and it concerns polymorphism. Usually we take lists to be polymorphic: for each type α there is a type α-`list`. However once we apply the `length` constructor to objects of any of these types we get objects of only the one type: `int`. We don't get a polymorphic family α-`int`, and nobody would normally suggest that we should. However, if one were an extreme purist one might note that, strictly speaking, Euler's totient function (for example) is properly defined only for those `int`s that are `int`s of lists of `int`s, not on `int`s that are `int`s of lists of `wombat`s, for example. However this purism is obviously extreme, since it's pretty clear that all these types are isomorphic and we will happily make do with only one type of `int`s. This will enable us to minute that fact that, for any natural number n, the set $[0, n-1]$ of its predecessors is of length n. Rosser called this the **Axiom of Counting**. The axiom of counting (for $\mathbb{N}$ at any rate) is fine[4]; it is the extension of this observation to ordinals that is ultimately problematic. The problem it ultimately leads to is the Burali-Forti paradox.

If we accept the transfinite version of Rosser's Axiom of Counting then the length of any initial segment A of the ordinals is the least ordinal α not in A. So what is the length of the (indisputably wellordered) collection On of all ordinals? It would have to be the least ordinal not in On! It seems that in order to avoid Burali-Forti one needs a stronger typing system that distinguishes between ordinals-from-(infinite)-lists-of-`a`s and ordinals-from-(infinite)-lists-of-`b`s. However the point at which hygiene compels one to adopt this stronger typing machinery comes a long way beyond ω_1. (It is not blindingly obvious that there are uncountable wellorderings—and we won't prove it—but in fact there are.) Fortunately for us we will be concerned in these notes with countable ordinals only; these dangers will remain innocuously over the horizon and we can quite safely take our ordinals to be monomorphic as we did our naturals, with the effect that we believe the analogue of the axiom of counting for countable ordinals.

Our concerns are fairly limited here and—mostly—this possibility of thinking of ordinals as relational types isn't important to us. However there are some ideas important to us for which this alternative concept of ordinals is indispensible, at least initially. Recall that a countable ordinal is an ordinal of a wellordering of a countable set. Since, by the Extended Axiom of Counting, every ordinal counts the set of its predecessors in the obvious ordering, this is the same as being an ordinal with countably many predecessors. The set of countable ordinals is sometimes called the **Second Number Class**. The countable ordinals are naturally ordered by magnitude, so the Second Number Class has an ordinal. By the Extended Axiom of Counting its ordinal must be the least ordinal not in it, and the least ordinal not in it is obviously the first uncountable ordinal. This ordinal is called 'ω_1'. If we ask not for the length of the set ordered by magnitude, but for its cardinal number ("How many countable ordinals are there?"), the answer is '$\aleph_1$'.

Second Number Class

We also need the concept of the **cofinality** of an ordinal.

Definition 10 *The cofinality $cf(\alpha)$ of an ordinal α is the least ordinal that is the length of an unbounded subset of a wellordering of length α.*

[4]There is virtue to be gained from thinking about how one might prove it!

At this point we should minute the standard fact that—assuming the countable axiom of choice—$cf(\omega_1) = \omega_1$. Suppose not, and let $\langle X, < \rangle$ be a wellordering of length ω_1 with $x_1 < x_2 < x_3 \ldots$ a cofinal subsequence of length ω. Then if we let $X_n =: \{x \in X : x_i \leq x < x_{i+1}\}$ then all the X_n are countable (ω_1 is the least uncountable ordinal after all) so $\{X_n : n \in \mathbb{N}\}$ is a partition of the uncountable set X into countably many countable pieces. Countable choice tells us that a union of countably many countable sets is countable.

(and yes, X here could be taken to be the second number class).

Notice that cf is **idempotent**: $cf(cf(\alpha)) = cf(\alpha)$. This is because "is a cofinal subsequence of" is transitive.

THEOREM 11 *Every countable limit ordinal is of cofinality ω.*

Proof:

Let α be a countable limit ordinal. Then there is a worder $<_\alpha$ of $\mathbb{N}$ of order type α. We now define a cofinal subsequence of $\langle \mathbb{N}, <_\alpha \rangle$ of length ω. The first point is 0. Thereafter then $n+1$th point is the smallest natural number which is $>_\alpha$ the nth point.

Why is this sequence cofinal? Suppose it reaches a limit below α. Consider a natural number above this limit. It must be below something (say the nth) in the list of points we have identified, since this list contains arbitrarily large natural numbers. But then, at that stage n, it was a better candidate to be the nth point than the point we chose. ■

It is very important that this construction of a cofinal sequence for α needs an extra input, namely the wellordering $<_\alpha$ of $\mathbb{N}$. If we vary the choice of α we get a different cofinal sequence.

There doesn't seem to be any uniform way of devising a cofinal sequence for α that works solely from the order structure of α itself. Of course if we had a uniform way of devising—on being given a countable ordinal β—a bijection between $\mathbb{N}$ and the ordinals below β, then we could plug that bijection into the construction of theorem 11 to obtain a uniform construction of a family of cofinal sequences. But then we don't have a uniform way of obtaining such bijections either! This fact is an easy consequence of a rather hard result, namely that ZF does not prove that $\aleph_1 \leq 2^{\aleph_0}$. *ZFC* proves that $\aleph_1 \leq 2^{\aleph_0}$, and ZF proves that $\aleph_1 \leq^* 2^{\aleph_0}$. If we had a uniform way of selecting, for each countable ordinal, a wellordering of $\mathbb{N}$ of that length, then we would have an injection from the second number class (which is of size $\aleph_1$) into the reals. (This will be remark 23.)

The picture of ordinals as order-types of wellorderings also gives us slightly smoother—and more fundamental—motivations for the operations of addition, multiplication and exponentiation of ordinals that we have already seen. Addition corresponds to disjoint union (concatenation) and multiplication to colex order of the product. It is worth noting that because these definitions do not involve recursion we can invoke them in connection with linear order types that are not wellfounded: they work for arbitrary total order types. And the operations obey the distributivity laws that you expect.

Another operation that is motivated by the picture of ordinals as isomorphism types of wellorderings is subtraction: when $\beta \leq \alpha$ we say $\alpha - \beta$ is the length of a well-ordering obtained by chopping off from a well-ordering of length α the unique initial segment of length β. Observe however that for subtraction of β from α to be well-defined we need (i) an ordering of type α to have a unique initial segment of type β—or at the very least we need (ii) all the tail segments that remain after deleting of an initial segment of type β to be isomorphic. Thus we can subtract ω^* from $\omega^* + \omega$ to obtain ω but to get subtraction of β from α to be defined for all $\beta \leq \alpha$ we need all initial segments of an ordering of type α to be pairwise nonisomorphic—or something quite like it—and there is not much hope of that unless α is an ordinal.

Exercise 2 *Give a recursive definition of ordinal subtraction, and prove that your definition obeys $\beta + (\alpha - \beta) = \alpha$.*

Ther is one other fact about ordinals we will need which can be obtained only from the ordinals-as-isomorphism-classes-of-wellorderings view. Here we will be concerned specifically with countable ordinals. Recall that a countable ordinal α is the length of a wellordering of a countable set. So without loss of generality α is the length of a wellordering of $\mathbb{N}$. A wellordering of $\mathbb{N}$ can be coded as a set of ordered pairs of naturals, and ordered pairs of naturals can be coded as naturals. Wellorderings of $\mathbb{N}$ can therefore be coded as sets of naturals, which is to say as *reals*. This means that there is a surjection from the set of reals to the set of countable ordinals as follows: if a real codes a wellordering of $\mathbb{N}$, send it to its length, `else` 0. Notice that this does not obviously give us an injection from the second number class into the reals: to do that we would have to choose, for each countable ordinal, a wellordering of the naturals of that length, and there is no obvious way to choose one. Notice that countable choice does not help here. We shall see more of this later.

A countable ordinal is an ordinal that is the length of a wellordering of $\mathbb{N}$ or of a subset of $\mathbb{N}$—it makes no difference. Cantor called the set of countably ordinals the *Second Number Class* (the first number class is $\mathbb{N}$). A *recursive* ordinal is an ordinal that is the length of a *recursive* wellordering of $\mathbb{N}$ or of a subset of $\mathbb{N}$—it makes no difference, that is to say, a wellordering whose graph (set of ordered pairs of natural numbers) is a recursive (= decidable) set. A decidable relation on an infinite subset of $\mathbb{N}$ is isomorphic to a decidable relation on the whole of $\mathbb{N}$ because the function enumerating the decidable subset is itself decidable.

There is a simple cardinality argument to the effect that not every countable ordinal is recursive, and I will go over it in some detail in lectures. Rosser's extended axiom of counting tells us that the length of the wellordering of all the countable ordinals has uncountable length, so there are uncountably many (in fact $\aleph_1$) countable ordinals. However the set of recursive ordinals is a surjective image of the set of all machines, and that set is countable. Clearly every recursive ordinal is countable, so there must be countable ordinals that are not recursive.

DEFINITION 12
The least nonrecursive ordinal is the **Church-Kleene** ω_1*, aka* ω_1^{CK}.

A standard application of countable choice tells us that every countable set of countable ordinals is bounded below ω_1, so we know that ω_1^{CK} is actually a countable ordinal. But we can do much better than that, and without using the axiom of choice.

REMARK 13 *The family of recursive ordinals is a proper initial segment of the second number class.*

Proof:
Suppose $<_R$ is a wellordering of $\mathbb{N}$ whose graph is a decidable subset of $\mathbb{N} \times \mathbb{N}$. That is to say that the length of $<_R$ is a recursive ordinal. Now consider any ordinal α less than the length of R. This is the length of a proper initial segment of $<_R$—of $<_R \restriction \{m \in \mathbb{N} : m <_R n\}$ for some n, say—and this initial segment of $<_R$ is a decidable subset of $\mathbb{N} \times \mathbb{N}$ (it has the number n as a parameter) and its length is therefore a recursive ordinal. ■

REMARK 14 *Every recursive limit ordinal has cofinality* ω*—recursively. That is to say: whenever* R *is a decidable binary relation on* $\mathbb{N}$ *that wellorders* $\mathbb{N}$ *to a length that is a limit ordinal there is* $X \subseteq \mathbb{N}$ *s.t.* $otp(R \restriction X) = \omega$.

Proof: The usual proof works. We enumerate the members of X in increasing order $x_0, x_1 \ldots$. We set $x_0 := 0$. Thereafter x_{n+1} is the least natural number x such that $\langle x_n, x \rangle \in R$. This is clearly an effective procedure. ■

EXERCISE 3 *The class of recursive ordinals is closed under the Doner-Tarski function* f_α *for every recursive ordinal* α.

Something to be alert to. Do not confuse the concept of a recursive ordinal with the concept of a *recursive pseudowellordering* of $\mathbb{N}$. This would be a binary relation R on $\mathbb{N}$ which is total orders with the property that every decidable subset of $\mathbb{N}$ has an R-least member.

When reasoning inside a formal system of arithmetic are is needed in approaching the concept of recursive ordinal. It's one thing to have a binary relation on $\mathbb{N}$, it is quite another to have a proof that this binary relation is a wellorder. Come to think of it, how on earth can a system of first-order arithmetic (such as Peano Arithmetic) *ever* prove that a binary relation is wellfounded? After all, to show that a relation is wellfounded one has to be able to reason about all the subsets of its domain, and a first-order theory cannot reason about arbitrary subsets. The answer is that whenever T (being a first order theory of arithmetic) proves that a relation R on $\mathbb{N}$ is a wellorder what is going on is that T proves all instances of R-recursion that can be expressed in the language of T.

1.4 Normal functions

As usual, a set is closed iff it contains all its limit points.

Brief chat here about the order topology.

Definition 15

*A **clubset** is a **CL**osed and **U**n**B**ounded set, or, alternatively, the range of a total continuous function.*

*A **normal** function f is*
(i) continuous: $f(\mathtt{sup}(A)) = \mathtt{sup}(f``A)$; and
(ii) strictly increasing: $\alpha < \beta \to f(\alpha) < f(\beta)$.

A clubset might be the range of lots of distinct continuous functions (repetitions are allowed, after all), but it is the range of only one *normal* function, to wit: the function that enumerates it. This bijection between the class of clubsets and the class of normal functions will be very useful to us. At times it will almost feel as if we have a datatype whose members can be thought of as clubsets and as normal functions at will. (cf page **??**)

Clearly the derived set of a clubset is club.

It is easy to show that every normal function has a fixed point. If f is normal, then $\mathtt{sup}\{f^n\alpha : n \in \mathbb{N}\}$ is the least fixed point for f above α. In fact:

Lemma 16 *The function enumerating the set of fixed points of a normal function is also normal.*

Proof:
This needs only the observation that if f is continuous then the sup of any set of fixed points for f is also fixed. ∎

Notice that we can now define, in a completely straightforward way, a transfinite sequence of normal functions from the second number class into itself—or, indeed, from the class of all ordinals into itself. Let C_0 be the set of limit ordinals in the second number class, and $C_{\alpha+1}$ be the limit points of C_α. We say $C_{\alpha+1}$ is the **derived set** [from] C_α. Take intersections at limits. The sequence of C_αs is the sequence of derived sets. Now let f_α be the function that enumerates C_α. C_α is club so f_α is normal.

(It might be an idea to think about what these functions actually are.)

If f is a a normal function then for any ordinal α we can define a function f^α as follows:

$f^0(\beta) =? \ldots 1$, presumably

$$f^1(\beta) = f(\beta)$$

$$f^{\alpha+1}(\beta) = f(f^\alpha(\beta))$$

$$f^\lambda(\beta) = \mathtt{sup}\{f^\zeta(\beta) : \zeta < \lambda\}$$

Must check that f^α is normal if f is.

NB sez: if f = succ then $f^\omega(\omega) = \omega$ so f^ω is not strictly increasing and isn't normal; tf sez: this is because succ is not normal!

There is another way of defining unary functions from ordinals to ordinals that gives us—I was about to say *functions that are more rapidly increasing*. That wouldn't be quite correct: every function that is given by the second method is also given by the first method, but the first method takes longer to reach it.

The second method defines C_0 to be the set of limit ordinals in the second number class as before. We take intersections at limits as before. As before f_α will be the normal function that enumerates C_α. However now $C_{\alpha+1}$ is defined to be the set of fixed points of the normal function f_α that enumerates C_α.

(Notice that although the first method could have started with $C_0 =$: second number class, the second method can't.)

We should think a bit here about what this new series of increasing functions look like.

There is a difference between these two ways of getting fast-growing functions that may strike a chord with people used to type disciplines. In the first case we can think of the ordinals that are arguments and the ordinals that are values as being two different types: green ordinals and blue ordinals.

In both cases we are indexing, by the ordinals, a family of ever-shrinking subsets of the ordinals. (We identify each skinny subset with the function that enumerates it). In both caes we take intersections at limits. In the first case the next set after A is the collection of limit points of A: we define a function from ordinals to sets of reals. $f(\alpha + 1)$ is the set of limit points of $f(\alpha)$. Nothing in this first construction compels us to think of the shrinking sets as shrinking sets *of ordinals*. Indeed in Cantor's original setting the derived sets are all sets of *reals* not sets of ordinals.

The difference between the first and second methods lies in the successor step. In the first construction the derived set at each stage is constrained to be a subset of the set it was derived from, so its members are objects of the same flavour that were found in that set, and that is the only constraint on their nature. However the second method exploits fixed points, so the functions it speaks of must have its arguments and its values of the same type. That means that the derived set must be a set of ordinals.

The first method is strongly typed and produces functions that don't grow very fast. The second method produces fast-growing functions much more efficiently but is less strongly typed. What we are seeing here is another instance of the way in which relaxation of typing disciplines makes for greater strength.

1.5 Binary functions

Unary functions are all very well, but what we are interested in is binary functions. One thinks immediately of $+$ and $\cdot$ and exponentiation. The anisotropy of the ordinals makes these functions continuous in one variable but not in the other. In fact they are *normal* in one variable but not in the other. For any given α the function $\beta \mapsto \alpha + \beta$ is normal but $\beta \mapsto \beta + \alpha$ is not; ditto $\beta \mapsto \alpha \cdot \beta$ and $\beta \mapsto \alpha^\beta$. The anisotropy means we have to be careful how we use them.

Doner-Tarski [4] consider a hierarchy of functions defined so that:

Definition 17

1. $f_0(\alpha, \gamma) =: \alpha + \gamma$;
2. $f_{n+1}(\alpha, 0) =: \alpha$;[5] check this!
3. $f_{n+1}(\alpha, \gamma + 1) = f_n(f_{n+1}(\alpha, \gamma), \alpha)$;
4. $f_{n+1}(\alpha, \lambda) =: \mathtt{sup}_{\gamma<\lambda} f_{n+1}(\alpha, \gamma)$;
5. $f_\lambda(\alpha, \beta) =: \mathtt{sup}_{\zeta<\lambda} f_\zeta(\alpha, \beta)$.

[Beware: we are already using the notation f_α in connection with C_α. However that was a nonce notation and should cause no confusion.]

If you forget this definition you will be able to reconstruct it quite easily by remembering two special cases: the recursive way in which multiplication is defined over addition and similarly exponentiation defined over multiplication. Indeed it is exactly that recursion which motivated the definition of this hierarchy in the first place.

Notice that all these functions are normal in their second argument. (One thing I would like to understand better is why the first few functions in this sequence correspond to natural operations on wellorderings but later ones don't. Part of the explanation must be that—probably from as early as two stages after exponentiation—the corresponding set theoretic operations involve transfinite iteration of power set, but it would be nice to be able to say something more intelligent than that.)

The functions are *not* continuous in their first argument. This apparently unremarkable fact explains something that would otherwise look a bit odd. The Doner-Tarski family arises by iterating an operation on functions $(On \times On) \to On$. Why isn't f_ω a fixed point for this operation? Iterating ω times is usually enough to get a fixed point .... Well either the operation is not suitably continuous, or the set of functions $(On \times On) \to On$ that we are considering is not a chain-complete poset. Presumably it is not a chain-complete poset, and this is something to do with them not being cts in the first argument. Notice, too, how the declaration of the Ackermann function (which incidentally is in $O(f_\omega)$! where this time f_ω is the ωth function in the fast-growing hierarchy in section 2 below) looks suspiciously like an announcement that it is a fixed point for the Doner-Tarski operation!

I would like to know how far you have to go along the Doner-Tarski hierarchy to get anything that dominates Ackermann. I don't suppose it matters much but i would sleep easier if i tho'rt i understood it.

1.6 Cantor's Normal Form Theorem

To prove Cantor's normal form theorem we will need to make frequent use of the following important triviality.

[5] This is surely correct. $f_{n+1}(\alpha, 0)$ must be the result of doing f_n of something or other 0 times to α and this must be α. The consideration that causes me slight unease is that according to this line of thought $\alpha \cdot 0$ should be α not 0. So the function we call multiplication—$\alpha \cdot \beta$ is actually $f_1(\alpha, \beta + 1)$. Not that it matters. But one would have expected to see something about this in the literature.

Remark 18 *If $f : On \to On$ is normal, then for every $\beta \in On$ there is a maximal $\alpha \in On$ such that $f(\alpha) \leq \beta$.*

Proof: Let α_0 be $\texttt{sup}\{\alpha : f(\alpha) \leq \beta\}$. By continuity of f

$$f(\alpha_0) = f(\texttt{sup}\{\alpha : f(\alpha) \leq \beta\})$$

which by continuity of f is

$$\texttt{sup}\{f(\alpha) : f(\alpha) \leq \beta\})$$

which of course is $\leq \beta$ since the ordinals are totally ordered. So α_0 is the largest element of $\{f(\alpha) : f(\alpha) \leq \beta\})$. ■

The way into Cantor Normal Forms is to think of remark 17 as a rudimentary result of the kind "Given an ordinal β and a normal function f, $f(\alpha_0)$ is the best approximation to β from below that I can give using f." Cantor Normal form is an elaboration of this idea into a technique. Let us first minute a few normal functions to see what sort of things we can attack β with. For every $\alpha > 0$ the functions

$$\gamma \mapsto \alpha + \gamma; \qquad \gamma \mapsto \alpha \cdot \gamma; \qquad \gamma \mapsto \alpha^{\gamma}$$

are all normal, and each is obtained by iteration from the preceding one.

We are given β and we want to express it in terms of a normal function. Let α be some random ordinal below β. Then $\gamma \mapsto \alpha^{\gamma}$ is a normal function and since $\alpha < \beta$ we know by remark 17 that there is a largest γ such that $\alpha^{\gamma} \leq \beta$. Call this ordinal γ_0. Then $\alpha^{\gamma_0} \leq \beta$. If $\alpha^{\gamma_0} = \beta$ we stop there.

Now consider the case where $\alpha^{\gamma_0} < \beta$. By maximality of γ_0 we have

$$\alpha^{\gamma_0} < \beta < \alpha^{\gamma_0 + 1} = \alpha^{\gamma_0} \cdot \alpha \tag{*}$$

We now attack β again, but this time not with the normal function $\gamma \mapsto \alpha^{\gamma}$ but the function $\theta \mapsto \alpha^{\gamma_0} \cdot \theta$. So by remark 17 there is a maximal θ such that $\alpha^{\gamma_0} \cdot \theta \leq \beta$. Call it θ_0. By (*) we must have $\theta_0 < \alpha$.

If $\alpha^{\gamma_0} \cdot \theta_0 = \beta$ we stop there, so suppose $\alpha^{\gamma_0} \cdot \theta_0 < \beta$, and in fact

$$\alpha^{\gamma_0} \cdot \theta_0 < \beta < \alpha^{\gamma_0} \cdot (\theta_0 + 1) = \alpha^{\gamma_0} \cdot \theta_0 + \alpha^{\gamma_0} \tag{**}$$

by maximality of θ_0.

Now $\beta = \alpha^{\gamma_0} \cdot \theta_0 + \delta_0$ for some δ_0, and we know $\delta_0 < \alpha^{\gamma_0}$ because of (**).

What we have proved is that, given ordinals $\alpha < \beta$, we can express β as $\alpha^{\gamma_0} \cdot \theta_0 + \delta_0$ with γ_0 and θ_0 maximal. If $\delta_0 < \alpha$ we stop. However if $\delta_0 > \alpha$ we continue, by attacking δ_0 with the normal function $\gamma \mapsto \alpha^{\gamma}$.

What happens if we do this? We then have $\delta = \alpha^{\gamma_1} \cdot \theta_1 + \delta_1$, which is to say

$$\beta = \alpha^{\gamma_0} \cdot \theta_0 + \alpha^{\gamma_1} \cdot \theta_1 + \delta_1$$

One thing we can be sure of is that $\gamma_0 > \gamma_1$. This follows from the maximality of θ_0.

We now go back and repeat the process, this time with δ_1 and α rather than β and α.

Therefore, when we repeat the process to obtain:

$$\beta = \alpha^{\gamma_0} \cdot \theta_0 + \alpha^{\gamma_1} \cdot \theta_1 + \alpha^{\gamma_2} \cdot \theta_2 + \delta_3$$

and so on:

$$\beta = \alpha^{\gamma_0} \cdot \theta_0 + \alpha^{\gamma_1} \cdot \theta_1 + \alpha^{\gamma_2} \cdot \theta_2 + \ldots \alpha^{\gamma_n} \cdot \theta_n + \ldots$$

Now we do know that this process must terminate, because the sequence of ordinals $\{\gamma_0 > \gamma_1 > \gamma_2 > \ldots \gamma_n \ldots\}$ is a descending sequence of ordinals and must be finite, because $<_{On}$ is wellfounded.

So we have proved this:

THEOREM 19 *For all α and β there are $\gamma_0 > \ldots > \gamma_n$ and $\theta_0 \ldots \theta_n$ with $\theta_i < \alpha$ for each i, such that*

$$\beta = \alpha^{\gamma_0} \cdot \theta_0 + \alpha^{\gamma_1} \cdot \theta_1 + \alpha^{\gamma_2} \cdot \theta_2 + \ldots \alpha^{\gamma_n} \cdot \theta_n$$

■

In particular, if $\alpha = \omega$ all the θ_i are finite. Since every finite ordinal is a sum $1 + 1 + 1 + \ldots$ this means that every ordinal is a sum of a decreasing finite sequence of powers of ω.

Quite how useful this fact is when dealing with an arbitrary ordinal β will depend on β. After all, if $\beta = \omega^\beta$ then—if we run the algorithm with ω and β—all Cantor's normal form theorem will tell us is that this is, indeed, the case. Ordinals β s.t. $\beta = \omega^\beta$ are around in plenty. They are called *ϵ-numbers*. They are moderately important because if β is an ϵ-number then the ordinals below β are closed under exponentiation. The smallest ϵ-number is called 'ϵ_0'. For the moment what concerns us about ϵ_0 is that if we look at the proof of Cantor's Normal Form theorem in the case where β is an ordinal below ϵ_0 and $\alpha = \omega$ the result is something sensible. This is because, ϵ_0 being the *least* fixed point of $\alpha \mapsto \omega^\alpha$, if we apply the technique of remark 17 to some $\alpha < \epsilon_0$ the output of this process must be an expression containing ordinals below α.

Say something about doing it recursively to the exponents

Now we must ask a very mathematical question, one that might have occurred to you already. On what features of multiplication, exponentiation and addition does this construction actually rely? Suppose we have a family $\langle f_i : i \in On \rangle$ of functions of two arguments defined in the manner of definition 16 so that

$$f_{n+1}(\alpha, \gamma + 1) = f_n(f_{n+1}(\alpha, \gamma), \alpha). \tag{1}$$

(and we require $\gamma \mapsto f_{n+1}(\alpha, \gamma)$ to be continuous at limit γ. We'll worry later about what to do when the subscript is limit!).

Suppose we want to express a given β in terms of a given α and n. What do we need? We want the various f_n to be normal in at least one argument. That is to say, for each n and every ζ, the function $\tau \mapsto f_n(\zeta, \tau)$ must be normal. That way we can be sure—to return to our given β, α and n—that there is a *last* γ so that

$$f_n(\alpha, \gamma) \leq \beta$$

which is to say, there is a γ so that

$$f_n(\alpha, \gamma) \leq \beta < f_n(\alpha, \gamma + 1)$$

Of course if $f_n(\alpha, \gamma) = \beta$ we stop. Otherwise we have

$$f_n(\alpha, \gamma) < \beta < f_n(\alpha, \gamma + 1) = f_{n-1}(f_n(\alpha, \gamma), \alpha)$$

Now, by normality of $\zeta \mapsto f_{n-1}((f_n(\alpha, \gamma), \zeta)$, there will be a last δ such that

$$f_{n-1}((f_n(\alpha, \gamma), \delta) \leq \beta$$

and we repeat the process.

Notice that addition, multiplication and exponentiation are related as successive members of precisely this kind of sequence of functions:

$$f_0(\alpha, \beta) := \alpha + 1$$
$$f_1(\alpha, \beta) := \alpha + \beta$$
$$f_2(\alpha, \beta) := \alpha \cdot \beta$$
$$f_3(\alpha, \beta) := \alpha^\beta$$

So the definitions from definition 16 give rise to a system of ordinal notations.

Exercise 4 *Use Cantor Normal Forms to show that every ordinal can be expressed as a sum of powers of 2.*

The first nontrivial result we saw that involved exponentiation was Cantor's normal form theorem, theorem 18. It made us think about ordinals like ω^n and ω^ω. It would be nice to have natural examples of well-orderings of lengths other than ω. $\mathbb{N} \times \mathbb{N}$ ordered lexicographically is of length ω^2. And, in general, $\mathbb{N}^n$ ordered lexicographically is of length ω^n. We can well-order the set of all finite lists of natural numbers to a longer length than this by a variant of the lexicographic ordering, but the definition is forgettable because of complications that have to do with deciding how to compare lists of different lengths. In some ways a simpler way to present these ordinals is through well-orderings of polynomials by dominance.

Definition 20 *f **dominates** g if, for all sufficiently large n, $f(n) > g(n)$.*

Consider the quadratics $x \mapsto (ax^2 + bx + c)$ and order them by dominance. It is fairly clear that $x \mapsto (ax^2+bx+c)$ is dominated by $x \mapsto (a'x^2+b'x+c')$ iff $\langle a, b, c\rangle$ comes below $\langle a', b', c'\rangle$ in the lexicographic order of $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$. So the set of quadratics, ordered by dominance, is of length ω^3. In fact, the analogue of this holds for polynomials of higher degree as well: the set of polynomials of degree n, ordered by dominance, is of length ω^{n+1}. Finally, the set of all polynomials (ordered by dominance) will be of order $\omega + \omega^2 + \omega^3 \cdots + \omega^n \cdots$. What is this ordinal? Since $1+\omega = \omega$ it follows that $1+\omega$ copies of anything is the same length as ω copies of whatever it was, so in particular $\omega^n + \omega^{n+1} = \omega^n$. Given this, the sum is simply the sup of all these ordinals, which, by definition, is ω^ω. Of course we could have got straight the definition of the well-ordering of finite sequences of natural numbers for another presentation of ω^ω, but the advantage of this version is that it can be easily upgraded. Let us call this family of polynomials in one variable the set of polynomials of rank 1 (to give it a name). Now consider the set of polynomials in one variable with coefficients in $\mathbb{N}$ whose exponents are polynomials of rank 1. An example would be

Could say a bit more about this

$$x^{x^3+x} + x^{200} + 137 \cdot x^3.$$

These will be the polynomials of rank 2. If you order these by dominance you obtain a wellorder of length ω^{ω^ω}. Similarly an example of a poly of rank 3 would be

$$x^{x^x+x^3+x} + x^{x^{50}} + x^{200} + 137 \cdot x^3.$$

If we wellorder by dominance the set of all polynomials in one variable of finite rank we find it is of length ϵ_0.

1.7 The Veblen Hierarchy

The following old tripos question (which had an afterlife on PTJ's example sheet 4 for Part II Set theory and Logic) can be profitably reviewed here.

Exercise 5 *(Tripos IIA 1995 Paper 4 question 8, modified).*

Let $\mathcal{P} = \langle P, \leq\rangle$ and $\mathcal{Q} = \langle Q, \leq\rangle$ be chain-complete posets with least elements, and let $h : \mathcal{P} \times \mathcal{Q} \to \mathcal{P} \times \mathcal{Q}$ be a map which is order-preserving with respect to the pointwise product ordering. Let the two components of the ordered pair $h(x, y)$ be $h_1(x, y)$ and $h_2(x, y)$ respectively.

1. *Show that, for each fixed $x \in P$, the mapping $g_x : Q \to Q$ defined by $g_x(y) = h_2(x, y)$ is order-preserving. Let $m(x)$ be its least fixed point.*

2. *Show that the map $f : P \to P$ defined by $f(x) = h_1(x, m(x))$ is order-preserving. Let x_0 be its least fixed point.*

3. *Show that $\langle x_0, m(x_0)\rangle$ is the least fixed point of h.*

Stuff to fit in:

In this section we consider functions that enumerate fixed points. We obtain fixed points by iteration but we need regularity of ω_1 to make sure we don't disappear off the end.

Let **Veb** *of f be the function that enumerates the fixed points of f. What is* $\mathbf{Veb}^\lambda\ f\ (\zeta)$ *to be? Presumably* $Sup(\{\mathbf{Veb}^\alpha\ f\zeta : \alpha < \lambda\}$*. But be careful! The ζth fixed point of g could easily be less than the first fixed point of* **Veb** *g. Thus we could find that* $\mathbf{Veb}^\lambda\ f\ (\zeta)$ *could easily be equal to* $\mathbf{Veb}^\lambda\ f\ (\zeta+1)$*. We have to hit this function with* **Veb** *to obtain a normal function. Harold says we should consider what he calls* **fruitful** *functions. They lack the strictness condition.*

Let us return to lemma 15, and consider the more rapidly growing sequence of functions that it reveals to us. We can define a sequence of clubsets as follows. $P(0)$ is the set of ordinals of the form $\omega \cdot \alpha$, and[6] $\alpha \mapsto \phi(0, \alpha)$ is the function that enumerates $P(0)$. Thereafter $P(\alpha+1)$ is the set of fixed points of $\beta \mapsto \phi(\alpha, \beta)$, and $\beta \mapsto \phi(\alpha+1, \beta)$ (sometimes written '$\phi_{\alpha+1}$') is the function that enumerates $P_{\alpha+1}$. At limits $P(\lambda) =: \bigcap\{P(\zeta) : \zeta < \lambda\}$ and $\beta \mapsto \phi(\lambda, \beta)$ (sometimes written 'ϕ_λ') is the function that enumerates P_λ.

Note the following:

1. All the $P(\alpha)$ are club.
2. For each α there seems to be an ordinal P_α such that $P(\alpha) = \{P_\alpha \cdot \zeta : \zeta \in On\}$. If this is correct then the additive normal form alleged in the literature is easily explained!
3. $\{min(P(\alpha)) : \alpha \in On\}$ appears to be club;
4. Should we here be thinking about diagonal intersections...?

Now we can try to rerun the proof of the Cantor Normal Form theorem.

Suppose we are given an ordinal α, and we want a notation for it in terms of the ϕs. There will be a last ζ (call it ζ_0) such that $P(\zeta)$ contains an ordinal $\leq \alpha$. If $\phi(\zeta_0, 0)$ (which is the first member of $P(\zeta_0)$) is equal to α, we stop. If it isn't, we consider the normal function $\beta \mapsto \phi(\zeta_0, \beta)$. By normality there is a last β (call it β_0) such that $\phi(\zeta_0, \beta_0) \leq \alpha$. If $\phi(\zeta_0, \beta_0) = \alpha$ we stop. Otherwise we have $\phi(\zeta_0, \beta_0) < \alpha < \phi(\zeta_0, \beta_0 + 1)$.

How do we get from $\phi(\zeta_0, \beta_0)$ to α. We know that α is below the next fixed point above $\phi(\zeta_0, \beta_0)$, which is of course $\phi(\zeta_0, \beta_0+1)$. But note that $\phi(\zeta_0, \beta_0+1)$ is the sup of ω iterates of

more work needed here

[should perhaps bear in mind that the notion of critical ordinal is tied specfically to exponentiation rather than any other of the functions that are normal in one variable.]

Young Peter sez:

Given $\gamma \in$ On

$\exists$ last α_0 $\phi_{\alpha_0}(1) \leq \gamma < \phi_{\alpha_0+1}(1)$

[6] I think in the literature we start with exponents of ω but a slow start makes it easier to see what is going on!

and

$\exists$ last β_0 $\phi_{\alpha_0}(\beta_0) \leq \gamma < \phi_{\alpha_0}(\beta_0 + 1)$

So $(\forall \alpha' < \alpha)(\phi_{\alpha_0}(\beta_0) \in P_{\alpha'})$

So its a fixed point! In particular

$(\forall \alpha' < \alpha)(\phi_{\alpha'}(\phi_{\alpha_0}(\beta_0)) = \phi_{\alpha_0}(\beta_0))$

So

$\exists$ last α_1 $\phi_{\alpha_1}(\phi_{\alpha_0}((\beta_0) + 1)) \leq \gamma < \phi_{\alpha_1+1}(\phi_{\alpha_0}((\beta_0) + 1))$

and

$\exists$ last β_1 $\phi_{\alpha_1}(\phi_{\alpha_0}((\beta_0) + \beta_1)) \leq \gamma < \phi_{\alpha_1+1}(\phi_{\alpha_0}((\beta_0) + \beta_1 + 1))$

I suspect the theorem in Veblen must be:

$(\forall \alpha < \Gamma_0)(\exists \beta_1 \ldots \beta_n < \alpha)(\exists \gamma_1 \ldots \gamma_n < \alpha)(\alpha = \phi(\beta_1, \gamma_1) + \phi(\beta_2, \gamma_)) + \cdots \phi(\beta_n, \gamma_n))$

2 Fundamental sequences and fast-growing functions

My point of departure here is an exercise that my friend and colleague Peter Johnstone gives to the third-year logic students here at Cambridge: *prove that for every countable ordinal α there is a set of reals which is of order type α in the inherited order.* The students all try to do this by induction on α. This seems the obvious thing to do, but it is fraught with difficulties. Indeed it is precisely those difficulties which will be our concern here. So let me start by giving the proof that the students never think of, but which is in fact much easier.

Let α be an arbitrary countable linear order type (it doesn't even have to be an ordinal). Concatenate α copies of $\langle Q, 0 \rangle$ (the rationals as an ordered set with a designated constant.) This structure is a dense linear order with a family of designated constants forming a subset of order-type α. But the ordering formed by discarding the designated constants is a countable dense total order and is therfore isomorphic to the rationals. Therefore every countable linear order type embeds in the rationals. In particular, every countable ordinal embeds into the rationals and therefore into the reals.

[we can simplify this by using a "forth" construction]

So we know it can be done. Secure therefore in that knowledge, we can afford to try doing it the hard way.

We will show how to construct, for arbitrarily large countable ordinals α, an order-preserving map f_α from the ordinals below α into $[0, \infty)$—and i think we want the range of f_α to be unbounded whenever α is limit. Indeed we probably want the map to be continuous in the sense of the order topology on the ordinals. That is to say, if λ is a limit ordinal below α then $f_\alpha(\lambda)$ is the lub of f_α"$\{\beta : \beta < \lambda\}$. A nice enough construction will probably make this happen automatically.

The obvious candidate for f_ω is the function that sends the ordinal number n to the real number n. Thereafter we have two tricks we can use. If we have f_α we can construct $f_{\alpha\cdot\omega}$ by "squashing" the range of f_α down on $[0, 1)$ by composing with $\frac{2}{\pi}\tan^{-1}$ and then making copies to put in each interval $[n, n+1)$, and concatenating them. To be slightly less hand-wavy about it, let A_α be the range of f_α, then $f_{\alpha\cdot\omega}$ is the function that enumerates the points in

$$\bigcup_{n\in\mathbb{N}} \{n + (\frac{2}{\pi} tan^{-1})``A_\alpha\}$$

(where the notation '$n+X$' of course denotes $\{n+x : x \in X\}$.) If α is an ordinal that cannot be reached by this method we find an ω-sequence $\alpha_0 < \alpha_1 \ldots$ whose sup is α and compress the ranges of the f_{α_i} into the intervals $[i, i + 1)$, thus:

$$\bigcup_{n\in\mathbb{N}} \{n + (\frac{2}{\pi} tan^{-1})``A_{\alpha_n}\}$$

Let us suppose that we have such a family $\langle f_\alpha : \alpha < \omega_1\rangle$. Fix a countable ordinal ζ and consider the sequence $\langle f_\gamma(\zeta) : \gamma > \zeta\rangle$. It would be natural to expect this to be a non-increasing sequence of reals. After all, the more ordinals you squeeze into the domain of an f, the harder you have to press down on its values to fit all the arguments in. But you'd be wrong! Suppose that

$$(\forall\gamma < \gamma' < \omega_1)(\forall\zeta < \omega_1)(f_\gamma(\zeta) \geq f_{\gamma'}(\zeta)). \tag{2}$$

Then, for each $\zeta < \omega_1$, the sequence $\langle f_\gamma(\zeta) : \gamma > \zeta\rangle$ of values given to ζ must be eventually constant. For if it is *not* eventually constant then it has $cf(\omega_1) = \omega_1$ decrements, and we would have a sequence of reals of length ω_1^* in the inherited order, and this is known to be impossible.

So there is an eventually constant value given to ζ, which we shall write '$f_\infty(\zeta)$'. But now we have $\alpha < \beta \to f_\infty(\alpha) < f_\infty(\beta)$. (We really do have '$<$' not merely $\leq$' in the consequent: suppose $f_\infty(\alpha) = f_\infty(\beta)$ happened for some α and β; then for sufficiently large γ we would have $f_\gamma(\alpha) = f_\gamma(\beta)$ which is impossible because f_γ is injective). This means that f_∞ embeds the countable ordinals into $\mathbb{R}$ in an order-preserving way, and this is impossible for the same reasons.

So we conclude that the function $\langle\alpha, \beta\rangle \mapsto f_\alpha(\beta)$ is *not* reliably decreasing in its second argument.

So what can possibly have gone wrong? Surely any sensible allocation of maps to limit ordinals will be well-behaved in the sense that it obeys (2)? Let us step back a bit and introduce a new gadget, one which has been lurking in the background all along.

2.1 Fundamental sequences

Definition 21

A **fundamental sequence** *for a (countable) ordinal α is an ω-sequence of ordinals whose supremum is α.*

We will equivocate harmlessly between thinking of fundamental sequences as wellorderings and thinking of them as strictly increasing functions from $\mathbb{N}$ into the set of countable limit ordinals.

We will start by proving some facts about fundamental sequences.

Theorem 11 told us that every countable limit ordinal has cofinality ω. This is of course just the same as saying that every countable ordinal has a fundamental sequence.

Definition 22

A **family** $\mathcal{F}$ *is a function sending each limit ordinal in some given initial segment of the second number class to a fundamental sequence for that ordinal.*

How do we obtain families of fundamental sequences? Suppose the order type of the limit ordinals below α is successor, so $\alpha = \beta + \omega$. In those circumstances the obvious choice for a fundamental sequence for α is $\langle \beta + n : n < \omega \rangle$. So far so good. Now suppose in contrast that the limit ordinals below α form a sequence of order type β for some limit ordinal $\beta < \alpha$. That is to say, there is a function g from $\{\zeta : \zeta < \beta\}$ to the set of limit ordinals below α. But if there is also a fundamental sequence f for β, then $g \cdot f$ will be a fundamental sequence for α.

This last step works as long as the order type of the set of limit ordinals below α is less than α. If it isn't then one has to do something slightly more clever. If we consider the ordinals that are fixed points for the function that enumerates the limit ordinals—which is the problematic case we have just identified—what might this clever thing be? The function that enumerates the limit ordinals is $\alpha \mapsto \omega \cdot \alpha$. Let's keep our feet on the ground for the moment by considering its first fixed point, which is ω^ω. A fixed point $\geq \alpha$ for a normal function f can be obtained as sup $\{f^n(\alpha) : n \in \mathbb{N}\}$. So ω^ω is immediately presented to us as the sup of $\{\omega^n : n \in \mathbb{N}\}$ and this gives us a fundamental sequence for ω^ω.

The hope is that there will always be some generalisation of this construction however far out we go. If F is a normal function $On \to On$ then whenever $\langle \beta_n : n \in \mathbb{N} \rangle$ is a fundamental sequence for β then $\langle F(\beta_n) : n \in \mathbb{N} \rangle$ is a fundamental sequence for $F(\beta)$. We will return to this later.

We have just seen how the construction of a fundamental sequence for β needs as input a bijection between $\mathbb{N}$ and the set of ordinals below β. In fact we can refine the proof of theorem 11 by exhibiting an algorithm that takes a bijection between $\mathbb{N}$ and the ordinals below β (or takes a wellordering of $\mathbb{N}$ of length β) and returns a family of fundamental sequences for limit ordinals below β. Similarly there is an algorithm that takes a family of fundamental sequences for the ordinals below β and returns a bjiecton between $\mathbb{N}$ and the set of ordinals below β. (Really one should say that this algorithm accepts and outputs *notations* for these objects rather than the objects themselves. The notations are genuine finite objects and we can compute with them. A countable ordinal is not on the face of it a finite object: curiosity about how far one can go in thinking of countable ordinals as finite objects is the energy driving interest in the material in this tutorial.)

THEOREM 23
There is a natural map that takes a wellordering of $\mathbb{N}$ *of length* α *and returns a family of fundamental sequences for the limit ordinals below* α*—and vice versa.*

Proof:
This is a generalisation of theorem 11.

(i) *Left-to-right*

Suppose we have a wellordering $<_\alpha$ of the naturals to length α; let β be an arbitrary limit ordinal below α. We will find a sequence $\langle b_n : n \in \mathbb{N}\rangle$ of natural numbers which is of length ω according to $<_\alpha$, and whose sup in that order is the βth element of $\langle \mathbb{N}, <_\alpha\rangle$. We define b_0 to be the $<_{\mathbb{N}}$-least natural number in that unique initial segment of $\langle \mathbb{N}, <_\alpha\rangle$ that is of length β. Thereafter b_{n+1} is to be the $<_{\mathbb{N}}$-least natural number that belongs to that unique initial segment of $\langle \mathbb{N}, <_\alpha\rangle$ that is of length β and is $>_\alpha b_n$.

How do we know that the upper bound of this sequence is the βth element of $\langle \mathbb{N}, <_\alpha\rangle$? By construction the set $\{b_n : n \in \mathbb{N}\}$ is unbounded in $<_{\mathbb{N}}$. So if n is a natural number that lies above the $(\leq_\alpha)$-sup of $\{b_n : n \in \mathbb{N}\}$ but is still below the βth element then it is $<_{\mathbb{N}}$ terminally many of the b_n, and should have been chosen. Now we take β_n to be the length of the initial segment of $\langle \mathbb{N}, <_\alpha\rangle$ bounded by b_n.

Clearly we can do this simultaneously for all limit ordinals $\beta < \alpha$.

(ii) *Right-to-left*

We want to be able to construct a bijection between $\mathbb{N}$ and the ordinals below β on being given a family of fundamental sequences of limit ordinals below β.

The idea behind this proof is that the availability of fundamental sequences for limit ordinals below β enables us to give—in a uniform way—a finite description of any ordinal below β. Every infinite set of finite strings over a finite alphabet is demonstrably countable. Totally order the alphabet; then order the set of finite strings colex. It will be of length ω, as will any of its infinite subsets. So how do we get a finite notation for an arbitrary $\zeta < \beta$? Let $\{\beta_{0,n} : n \in \mathbb{N}\}$ be the fundamental sequence for β. Consider the first member of $\{\beta_{0,n} : n \in \mathbb{N}\}$ that is $\geq \zeta$. This is β_{0,n_0}, say. Record the n_0 `If` this β_n is actually equal to ζ `then HALT`, `else` step down from this ordinal to the last limit ordinal below it (which for the moment we will call 'α') and record the suffix 'i' such that it was $\beta_{0,i}$. (We don't need to record the decrement, and in any case if the fundamental sequence for α are sensible the α_i will be limit ordinals unless $\alpha = \omega \cdot (\gamma + n)$ for some $n < \omega$). Now let $\{\beta_{1,n} : n \in \mathbb{N}\}$ be the fundamental sequence for α. Consider the first member of $\{\beta_{1,n} : n \in \mathbb{N}\}$ that is $\geq \zeta$. `If` this $\beta_{1,n}$ is actually equal to ζ `then` record the n and `HALT`. `Else` step down from this ordinal to the last limit ordinal below it (which for the moment we will call 'α' as before) and record the suffix 'j' such that it was $\beta_{1,j}\ldots$ (As before we do not need to record the decrement). Eventually we will find ourselves a finite distance above a point of a fundamental sequence and this time we do record the decrement.

I think this has become garbled and should be rewritten

By this procedure we build a sequence of natural numbers. This sequence is going to have to be finite if this construction is to be of any use to us. The reason why it will be finite is that the sequence of ordinals that were named 'α' at any stage of this process form a strictly descending sequence of ordinals and so must be finite.

So we have coded every ordinal below β by a finite string of symbols, and thence—using standard methods—by a natural number.

Perhaps we should explain how, with the help of this notation for ζ, we can navigate our way thither from 0. Given a sequence s for ζ we recover ζ as follows. First approximation is $\beta_{s(1)}$. Step down to the last limit ordinal below $\beta_{s(1)}$. Second approximation is the $s(2)$th member of the fundamental sequence for the last limit ordinal below $\beta_{s(1)}$. The last member of s (that is, $s(|S|)$) tells us what natural number to subtract from the approximation-in-hand.

To do this we think of the family as a set of ordered pairs $\langle s, s' \rangle$ of these finite sequences where (the ordinal notated by s) $<$ (the ordinal notated by s'). ■

REMARK 24 *There is no definable family of fundamental sequences for all $\alpha < \omega_1$.*

Proof: Let $\mathcal{F}$ be a family of fundamental sequences for all countable limit ordinals. We will show that $\mathcal{F}$ cannot be definable.

We define by recursion on the second number class a sequence $\langle W_\alpha : \alpha < \omega_1 \rangle$ of wellorderings of $\mathbb{N}$ (so each is a subset of $\mathbb{N} \times \mathbb{N}$). We fix once for all a bijection $\mathbb{N} \times \mathbb{N} \longleftrightarrow \mathbb{N}$.

0 is easy; successor steps are easy; at a limit λ use the fundamental sequence $\mathcal{F}\lambda$, to get the codes $W_{\mathcal{F}\lambda n}$ you have already formed for each $\mathcal{F}\lambda n$ and then piece them all together one after the other to get a wellordering of $\mathbb{N} \times \mathbb{N}$. Use the bijection $\mathbb{N} \times \mathbb{N} \longleftrightarrow \mathbb{N}$ to turn this into a code for $\Sigma_{n \in \mathbb{N}} \mathcal{F}\lambda n$, which we will call λ'. Here we have to be careful, because the *sum* of a sequence of ordinals might be bigger than its *supremum*. What we want is a wellordering of $\mathbb{N}$ to the sup of this set of ordinals (which is λ) not its sum (which is λ'). Suppose $\lambda' > \lambda$. We delete from $\mathbb{N}$ those naturals that get sent to addresses after λ, and we delete ordered pairs containing them from the graph of the wellordering of $\mathbb{N}$ to length λ'. What's left is a wellordering of a proper subset $\mathbb{N}' \subset \mathbb{N}$ to length λ. But there is an obvious canonical bijection between $\mathbb{N}'$ and $\mathbb{N}$, and we can use it to copy the wellordering of $\mathbb{N}'$ over to a wellordering of $\mathbb{N}$ to length λ as desired. None of this uses any AC.

This shows that if we have a function $\mathcal{F}$ assigning a fundamental sequence to every countable limit ordinal, then we have a function assigning to each countable ordinal a wellordering of $\mathbb{N} \times \mathbb{N}$ of that length, and this new function can be defined in terms of $\mathcal{F}$. But (as we saw on page 11) any wellordering of $\mathbb{N} \times \mathbb{N}$ is coded by a real number so the existence of the new function assigning a fundamental sequence to every countable ordinal implies $\aleph_1 \leq 2^{\aleph_0}$. It is known that this is independent of ZF.

This doesn't mean that there can be no family $\mathcal{F}$ of fundamental sequences for all countable limit ordinals, but it does mean that no such family can be definable; if it were, we would have an outright proof that $\aleph_1 \leq 2^{\aleph_0}$.

■

2.2 Fast-growing hierarchies

Our motive for considering fundamental sequences is that any family of fundamental sequences can be used to extend declarations of families of functions $\mathbb{N} \to \mathbb{N}$ into the transfinite in something like the following style.

The first person to spell out a fast-growing hierarchy seems to have been Hardy [12]. His idea was that if you could extend a fast-growing hierarchy out to all countable ordinals then you would have an injection of the second number class into the reals. As we have just seen, this hope is vain.

DEFINITION 25 *Suppose $\mathcal{F}$ is a family in the sense of definition 20. Then we can declare*

$f_0^{\mathcal{F}}$ = *some function or other;*

$f_{\alpha+1}^{\mathcal{F}}$ =: *do something to* f_α*;*

$f_\lambda^{\mathcal{F}}(n) =: f_{(\mathcal{F}\ \lambda\ n)}(n)$.

(Typically we will omit the '$\mathcal{F}$' superscript).

There is also the (apparently) minor detail that in the process of constructing the embeddings f_α from initial segments of the second number class into the reals we exploit representations of countable ordinals as *sums* of countably many smaller ordinals whereas in the definition of the fast-growing hierarchies we exploit fundamental sequences—which are representations of limit ordinals as *suprema* of ω-sequences of small ordinals. I don't think the difference matters, but one never knows.

Declarations in the style of definition 24 are typically used to generate families of functions where f_α dominates f_β whenever $\beta < \alpha$.

At successor stages this will be taken care of by the second clause and the purpose of the third clause is to ensure that f_λ dominates ("majorises") f_β with $\beta < \lambda$ for λ limit. Naturally one expects that if f_0 was strictly increasing then all the later f_α will be too—and that one will be able to prove this by transfinite induction. However to arrange for strict monotonicity of all the f_α it turns out one needs a condition on the family $\mathcal{F}$ of fundamental sequences which we will now investigate.

[*HOLE Stuff to fit in*

Let $\mathcal{F}$ be a counted family of functions, equipped with $F : \mathbb{N} \to \mathcal{F}$. Then we can define a supremum f_∞ of $\mathcal{F}$ by

$$f_\infty(n) = sup\{(F\ i\ n) + 1 : i \leq n\}$$

We need this in the case where $\{\mathcal{F}\}$ is $\{f_\beta : \beta < \alpha\}$]

2.2.1 Schmidt Coherence

The idea is to prove by induction on α that f_α is monotone increasing and dominates all earlier f_β. Let's get the dominance out of the way first. Given the induction hypothesis it's easy to prove that f_α dominates all earlier f_β. The successor case is obvious; for the limit case suppose f_{λ_i} is strictly increasing for each $i \in \mathbb{N}$ and that later fs dominate earlier fs. If f_λ is $n \mapsto f_{\lambda_n}(n)$ then it dominates every f_{λ_i}.

Mark Ryten sez that Schmidt construx works only where $f_{\alpha+1}$ is pointwise bigger than f_α and guarantees pointwise for this!

NB denies this: says one needs Schmidt-coherence even for domination over Schmidt relatives.

Why isn't strict monotonicity obvious too? If f_α is strictly increasing so is $f_{\alpha+1}$. The hard case is that of limit ordinals. Let λ be limit and $\langle \lambda_n : n \in \mathbb{N} \rangle$ the fundamental sequence for it.

We want

$$f_\lambda(n) < f_\lambda(n+1). \qquad \text{This holds iff}$$

$$f_{\lambda_n}(n) < f_{\lambda_{n+1}}(n+1). \qquad \text{Now we do at least have}$$

$$f_{\lambda_n}(n) < f_{\lambda_n}(n+1)$$

because f_{λ_n} is strictly increasing by induction hypothesis. So to complete the proof it will suffice to show

$$f_{\lambda_n}(n+1) < f_{\lambda_{n+1}}(n+1)$$

which will follow if $(\forall\lambda\forall n)(\texttt{succ}(\lambda_n, \lambda_{n+1}))$ where $\texttt{succ}(\alpha, \beta)$ is:

$$\alpha < \beta \to (\forall m)(f_\alpha(m) < f_\beta(m)).$$

However when β is a limit we can be sure of the consequent of $\texttt{succ}(\alpha, \beta)$ only for sufficiently large m. The construction of the f_αs ensures that $\texttt{succ}(\alpha, \beta)$ holds if $\beta = \alpha + 1$ or if β is limit and $\alpha = \beta_0$. To be sure of $\texttt{succ}(\alpha, \beta)$ when $\alpha < \beta$ are members of a fundamental sequence we need to specify that they are related by the transitive closure of the union of these two relations. A family of fundamental sequences satisfying this condition is **Schmidt-coherent**.

Formally:

Definition 26 *Let the family $\mathcal{F} : \Delta \to \Delta^\omega$ be an assignment of fundamental sequences to an initial segment Δ of the second number class. Let $<_\mathcal{F}$ be the strict partial order which is the transitive closure of $\beta <_\mathcal{F} \beta+1$ and $(\mathcal{F}\ \beta)\ 0 <_\mathcal{F} \beta$. (Schmidt [16] calls $<_\mathcal{F}$ the step-down relation of $\mathcal{F}$.)*

Then

$\mathcal{F}$ is **Schmidt-coherent** *iff*

$$(\forall\lambda \in \Delta)(\lambda \textit{ limit} \to (\forall n \in \mathbb{N})((\mathcal{F}\ \lambda\ n) <_\mathcal{F} (\mathcal{F}\ \lambda\ (n+1)))).$$

(Schmidt calls these 'built-up' rather than 'coherent'.)

It is not hard to see that, for any $\mathcal{F}$, $<_\mathcal{F}$ is a wellfounded (upward-branching) tree and that all paths are of length ω. One steps down at limit ordinals λ by leaping downwards to $\mathcal{F}\ \lambda\ 0$—the first member of the fundamental sequence for λ, aka λ_0. At successor steps one subtracts one. The way one steps down is uniquely determined by where one **is** not by where one **starts from**. This means that two descending paths that meet anywhere thereafter remain coincident.

Schmidt-coherence is equivalent to the condition that every fundamental sequence lies entirely within one branch of the tree.

EXERCISE 6 *Define the natural assignment of fundamental sequences to ordinals below ϵ_0 and check that it is Schmidt-coherent.*
Do the same for the ordinals below Γ_0.

This completes the proof of:

THEOREM 27 *(Schmidt [16] theorem 1)*

If $\mathcal{F}$ is a Schmidt-coherent family of fundamental sequences then every function in the fast growing hierarchy over $\mathcal{F}$ is monotone and strictly increasing.

Proof: The definition of Schmidt-coherence was cooked up precisely to make this work.

■

LEMMA 28 *If $\mathcal{F}$ is Schmidt-coherent, λ is limit and $n \in \mathbb{N}$ then $\mathcal{F}^{\lambda,n}$, defined by*

$$\mathcal{F}^{\lambda,n}\ \lambda\ m =: \mathcal{F}\ \lambda\ (m+n);\quad \mathcal{F}^{\lambda,n}\ \beta\ m = \mathcal{F}\ \beta\ m \textit{ for other } \beta$$

...is also Schmidt-coherent.

Proof: It will suffice to show that $\mathcal{F}\ \lambda\ 0 <_{\mathcal{F}^{\beta,n}} \lambda$. But—since $\mathcal{F}$ is Schmidt-coherent we have $\mathcal{F}\ \lambda\ 0 <_{\mathcal{F}} \lambda$. Hence—by the definition of $\mathcal{F}^{\beta,n}$—we have $\mathcal{F}^{\beta,n}\ \lambda\ 0 <_{\mathcal{F}^{\beta,n}} \mathcal{F}^{\beta,n}\ \lambda\ n$. But this last ordinal is the $<_{\mathcal{F}^{\beta,n}}$-predecessor of λ, whence $\mathcal{F}\ \lambda\ 0 <_{\mathcal{F}} \mathcal{F}\ \lambda\ n <_{\mathcal{F}^{\beta,n}} \lambda$.

LEMMA 29 *Let $\mathcal{F}$ be a Schmidt-coherent system of fundamental sequences for Δ an initial segment of the second number class, and suppose $\alpha < \beta \in \Delta$. Then there is a system $\mathcal{F}^{(\alpha,\beta)}$ of fundamental sequences*[7] *for Δ such that*

1. *$\mathcal{F}^{(\alpha,\beta)}$ is Schmidt-coherent;*
2. *$\alpha <_{\mathcal{F}^{(\alpha,\beta)}} \beta$ and*
3. *for all $\delta \leq \alpha$ we have $\mathcal{F}^{(\alpha,\beta)}\delta = \mathcal{F}\delta$.*

Proof:
(lifted brazenly from Schmidt [16])
We define a sequence $\langle \gamma_n, \mathcal{F}_n \rangle$ as follows.
$\gamma_0 =: \beta$, $\mathcal{F}_0 =: \mathcal{F}$;
Thereafter

[7] This is my notation not hers, and i've put in the brackets to make it less likely that readers will confuse it with the "$\mathcal{F}^{\alpha,\beta}$"

- if $\gamma_n = \alpha$ then $\gamma_{n+1} =: \alpha$ too, and $\mathcal{F}_{n+1} =: \mathcal{F}_n$;
- if $\gamma_n = \delta + 1 > \alpha$ then $\gamma_{n+1} =: \delta$ and $\mathcal{F}_{n+1} =: \mathcal{F}_n$;
- if $\gamma_n > \alpha$ and is a limit, and m is minimal such that $\mathcal{F}\gamma_n m \geq \alpha$ then $\gamma_{n+1} =: \mathcal{F}\gamma_n m$ and
-
  - if $\gamma \neq \gamma_n$ then $\mathcal{F}_{n+1}\gamma q =: \mathcal{F}_n\gamma q$, and
  - if $\gamma = \gamma_n$ then $\mathcal{F}_{n+1}\gamma q =: \mathcal{F}_n\gamma(q + m)$.

Using lemma 27 it is easy to show that

- $\mathcal{F}_n$ is Schmidt-coherent,
- $\gamma_n <_{\mathcal{F}_n} \beta$ or $\gamma_n = \beta$,
- $\mathcal{F}_n\delta = \mathcal{F}\delta$ for all $\delta < \alpha$,
- $\gamma_n \geq \alpha$.

Now $\langle \gamma_n : n < \omega \rangle$ is a nonincreasing sequence, so is eventually constant, so there is $n_0 \in \mathbb{N}$ such that $\gamma_{n_0} = \alpha$. Set $\mathcal{F}^{(\alpha,\beta)} =: \mathcal{F}_{n_0}$. ■

Lemma 30 *Let $\mathcal{F}$ be a Schmidt-coherent system of fundamental sequences for Δ an initial segment of the second number class, and let λ be the smallest limit ordinal not in Δ. Then there is a Schmidt-coherent system $\mathcal{F}'$ of fundamental sequences for $\Delta \cup \{\lambda\}$.*

Proof: Let $\langle \lambda_n : n \in \mathbb{N} \rangle$ be a fundamental sequence for λ. We define a sequence $\langle \mathcal{F}_n : n \in \mathbb{N} \rangle$ by recursion as follows. $\mathcal{F}_0 =: \mathcal{F}$ and thereafter $\mathcal{F}_{n+1} =: (\mathcal{F}_n)^{(\lambda_n, \lambda_{n+1})}$ as in 28. Now—by that lemma (itemwise!)—for each $n \in \mathbb{N}$ we have

1. $\mathcal{F}_n$ is Schmidt-coherent;
2. $\lambda_n <_{\mathcal{F}_{n+1}} \lambda_{n+1}$;
3. $\mathcal{F}_n\delta = \mathcal{F}_{n+m}\delta$ for all $\delta \leq \lambda_n$ and $m \in \mathbb{N}$.

We can now set $\mathcal{F}'\beta$ to be

- $\langle \lambda_n : n \in \mathbb{N} \rangle$ if $\beta = \lambda$;
- $\mathcal{F}\beta$ if $\beta \leq \lambda_0$;
- $\mathcal{F}_{m+1}\beta$ if $\lambda_m < \beta \leq \lambda_{m+1}$.

$\mathcal{F}'$ obviously assigns fundamental sequences to everything in $\Delta \cup \{\lambda\}$.

■

Theorem 31 *(Schmidt [16] theorem 2)*

Every proper initial segment of the second number class admits a Schmidt-coherent family of fundamental sequences.

Proof:
We prove by induction on 'α' that the countable ordinals strictly below α admit a Schmidt-coherent family.

The successor case is easy: if α is a successor of a successor, the assertion follows from the induction hypothesis; if α is the successor of a limit it follows from lemma 29 and the induction hypothesis.

So consider the case where α is limit.

Let $\langle \alpha_n : n \in \mathbb{N} \rangle$ be a fundamental sequence for α, and for each $n \in \mathbb{N}$ set $\sigma_n =: \Sigma_{m<n}\alpha_m$. Clearly $\alpha \leq \mathtt{sup}(\{\sigma_n : n \in \mathbb{N}\})$.

By the induction hypothesis for each $n \in \mathbb{N}$ there is a Schmidt-coherent family $\mathcal{F}_n$ for the ordinals below $\alpha_n + 1$. We now define a family $\mathcal{F}$ as follows:

$\mathcal{F}\gamma\ m =$:

- 0 if γ is zero or a successor;
- $\sigma_n + (\mathcal{F}(\gamma - \sigma_n)m)$ otherwise, where n is maximal so that $\sigma_n < \gamma$.

Now for all μ and ν such that $\sigma_n < \mu \leq \sigma_{n+1}$ and $\sigma_n < \nu \leq \sigma_{n+1}$ we have $\mu <_{\mathcal{F}} \mu \longleftrightarrow (\mu - \sigma_n) <_{\mathcal{F}_n} (\nu - \sigma_n)$. Hence if γ is a limit ordinal and $\sigma_n < \gamma \leq \sigma_{n+1}$ then $\gamma - \sigma_n$ is also a limit, and since $\mathcal{F}_n$ is Schmidt-coherent we have $\mathcal{F}(\gamma - \sigma_n)m <_{\mathcal{F}_n} \mathcal{F}(\gamma - \sigma_n)(m + 1)$ for each $m \in \mathbb{N}$. Thus $\mathcal{F}\gamma m = \sigma_n + (\mathcal{F}(\gamma - \sigma_n)m) <_{\mathcal{F}} \sigma_n + (\mathcal{F}(\gamma - \sigma_n)(m + 1)) = \mathcal{F}\gamma(m + 1)$. So $\mathcal{F}$ is Schmidt-coherent.

■

Can we omit 'proper' from the statement of theorem 30? We proved it without any use of AC.

> Rose says that theorem 30 is best possible, and credits Bachmann: *Transfinite Zahlen* Springer, 1967. I'm sceptical about this because he also says that Schmidt, too, proves that it is best possible—and she doesn't!
>
> If it really is best possible, it's presumably because a Schmidt-coherent family for all countable ordinals would give us an embedding of ω_1 into the reals, or something like that. There can be long sequences ($\geq \omega_1$) of functions with each function dominating all earlier functions, but they don't increase as fast as Wainer-Buchholtz.

what is this Rose reference?

We've managed to get this far on generalities that do not depend on the precise declaration of the fast-growing hierarchy. The time has now come to be specific. Let $\mathcal{F}$ be a Schmidt-coherent family of fundamental sequences.

The following seems to be popular: (Buchholtz-Wainer[3] refer to it merely as 'the' fast-growing hierarchy!)

Definition 32 *(Buchholtz-Wainer)*
The Fast-Growing Hierarchy

$f_0(x) =: x + 1$;
$f_{\alpha+1}(x) =: {f_\alpha}^{x+1}(x)$;
$f_\lambda(x) = f_{(\mathcal{F}\ \lambda\ x)}(x)$.

The fast-growing hierarchy with finite subscripts is the **Grzegorczyk** *hierarchy.*

The **Hardy Hierarchy** *([12]) is:*

$H_0(x) =: x + 1$;
$H_{\alpha+1}(x) =: H_\alpha(x + 1)$;
$H_\lambda(n) = H_{(\mathcal{F}\ \lambda\ n)}(n)$.

Just to reassure myself that i am in familiar surroundings i shall prove

REMARK 33 *For $\alpha < \omega$, f_α is primitive recursive.*

Proof: Clearly true for $\alpha = 0$. Define `iter` g so that $\mathtt{iter}(g, n) : m \mapsto (g^n(m))$ by means of the following declaration:

$$\mathtt{iter}(f, 0)\ m =: m;\ \mathtt{iter}(f, (n + 1))\ m =: f(\mathtt{iter}(f, n)\ m)$$

we see that $\mathtt{iter}(g, n)$ is primitive recursive as long as g is. Then $f_{\alpha+1} : n \mapsto (\mathtt{iter}(f_\alpha, n + 1)\ n)$ is primitive recursive as long as f_α is.

■

EXERCISE 7 *Determine f_0, f_1 and f_2.*

EXERCISE 8 *(Computer Science Tripos 1991:5:10)*
Ackermann's function is defined as follows:

$$A(0, y) =: y + 1;\ A(x + 1, 0) =: A(x, 1);\ A(x + 1, y + 1) =: A(x, A(x + 1, y))$$

For each n define
$a_n(y) =: A(n, y)$.
Prove $(\forall y)(\forall n \in \mathbb{N})(a_{n+1}(y) = a_n^{y+1}(1))$.

Notice that $a_0(x) = f_0(x) = x + 1$.

Then by induction on the recursive datatype of primitive recursive functions we prove that every primitive recursive function is dominated by all sufficiently late a_n.

THEOREM 34 *For every primitive recursive function $f(\vec{x}, n)$ there is a constant c_f such that*

$$(\forall n \forall \vec{x})(f(\vec{x}, n) < A(c_f, max(n, \vec{x})))$$

(In slang, every primitive recursive function is in O(Ackermann).)

Exercise 9 *Complete the proof.*

Notice that there cannot be a converse. This is because of the silly reason that there are slowly growing functions that are inverses of rapidly growing ones, and are therefore equally hard to compute. Try the computer science tripos question 1994 paper 5 question 11 (at `http://www.cl.cam.ac.uk/tripos/t-ComputationTheory.html`)

Then $A(n, n)$ diagonalises the a_n the way f_ω diagonalises the f_n. So $A(n, n)$ is "at the same level" as f_ω. In fact if f is primitive recursive, then the c_f of theorem 33 is precisely the level of the fast-growing hierarchy that f belongs to (I think!).

3 Consistency strength measured by ordinals

> Any not conspicuously deficient set theory can of course prove the existence of transfinite numbers without end, but this does not mean getting them all. What is so characteristic of the transfinite is that we then go on iterating the iteration, iterating the iteration of the iterations, and so on, until somehow our apparatus buckles; and the least transfinite number after the buckling of the apparatus is how strong the apparatus was.
>
> W.V.Quine: [15] pp 323-4

Maybe we should say something here about how the endeavour to achieve a complete consistent system of arithmetic by transfinitely adding Gödel sentences comes unstuck. Quite where it comes unstuck will presumably depend on the strength of the original system. For PA it comes unstuck at ϵ_0?

4 Some illumination from Nathan Bowler

Suppose we have an ordinal x that we want to describe. We begin analysing it as for the Cantor normal form. So we pick a maximal y such that $\omega^y \leq x$. Then we pick a maximal n such that $\omega^y \cdot n \leq x$ and a maximal z such that $\omega^y \cdot n + z \leq x$. Now we can show in the usual way that:

- $n < \omega$;
- $z < x$;
- $x = \omega^y \cdot n + z$.

All of this is routine. The only thing we don't yet control is y. We know that $y \leq x$, and so we are only in trouble if $y = x$. The traditional thing to do at this point is to say "Our method has reached its limit. Let us give a name to the least ordinal at which this problem can occur, say ϵ_0, in order to get some closure."

However, we have some very strong information about the places where this trouble can happen. We must have $y = \omega^y$. That is, y is a fixed point of the map $\phi_0 : y \mapsto \omega^y$. So we simply enumerate the fixed points of ϕ_0 by some function ϕ_1 and get a new expression for y as $\phi_1(y')$, say. This will certainly be useful if we can guarantee that $y' < y$, but of course we can't.

But we can do the same trick again. If y is stroppy enough to equal y' then it must be ϕ_2 of something, where ϕ_2 enumerates the fixed points of ϕ_1. And we can keep going. At any stage, we simply enumerate the ordinals that have so far been badly behaved. So, for example, at a limit ordinal λ, we are in trouble iff $\phi_\alpha(y) = y$ for all $\alpha < \lambda$. So we choose ϕ_λ to enumerate the set of such y, which is the intersection of the fixed point sets of all the ϕ_α for $\alpha < \lambda$.

To summarise, we choose α maximal with the property that y is in the image of ϕ_α. Then we have $y = \phi_\alpha(y')$ with $y' < y \leq x$. Now we control everything. Except α. We are now in trouble if $\alpha = y$. So we say "Our method has reached its limit. Let us give a name to the least ordinal at which this problem can occur, say Γ_0, in order to get some closure."

Lemma 35 *For any x less than Γ_0 there exist n, α_1, α_2, $\dots \alpha_n$ and β_1, β_2, $\dots \beta_n$ with:*

- *$\phi_{\alpha_i}(\beta_i) \leq \phi_{\alpha_j}(\beta_j)$ for $i \leq j$;*
- *$\alpha_i < \alpha_{i+1}$ and $\beta_i < \alpha_i$ (for all i);*
- *$x = \phi_{\alpha_n}(\beta_n) + \phi_{\alpha_{n-1}}(\beta_{n-1}) + \dots + \phi_{\alpha_1}(\beta_1)$.*

Proof: : By contradiction. Let x be minimal with no such representation. I demonstrated above that we may find α, β, m and z with:

- $\alpha, \beta, z < x$;
- $m < \omega$;
- $x = \phi_\alpha(\beta) \cdot m + z$.

Now by minimality of x, we have a representation

$$z = \phi_{\alpha_n}(\beta_n) + \phi_{\alpha_{n-1}}(\beta_{n-1}) + \dots + \phi_{\alpha_1}(\beta_1).$$

So let $n' = n + m$, and let $\alpha_i = \alpha$ and $\beta_i = \beta$ for $n < i \leq n'$. We now have a representation of x in the desired form.

Now, given two such expressions constructed in this canonical way, (that is, so that no term of the form $\phi_\alpha(\beta)$ occurs with $\phi_\alpha(\beta) = \beta$, we can easily say which is the largest by the simple expedient of comparing corresponding terms in the two expressions.

Note that the above argument is precisely the same as that used for the Cantor normal form.

Further correspondence

tf:

> On the subject of Γ_0: i'm still not happy. What is the normal form for ω^{ϵ_0+2}?

NB:

> $\phi_0(\phi_{\phi_0(0)}(0) + \phi_0(0) + \phi_0(0))$

tf:

> I suppose i should now try to work out why. How does one get that from the CNF algorithm?

NB:

> I have written the derivation out in painful detail. If you fight your way through it then you should see how the extended algorithm works.

Step 1 : Find the maximal y such that $\omega^y \leq \omega^{\epsilon_0+2}$, the maximal n such that $\omega^y \cdot n \leq \omega^{\epsilon_0+2}$ and the maximal z such that $\omega^y \cdot n + z \leq \omega^{\epsilon_0+2}$. This is fairly easy as we are given x in the form ω^y. So we have $y = \epsilon_0 + 2$, $n = 1$ and $z = 0$. The expansion so far is ω^{ϵ_0+2}. Not very enlightening yet.

Step 2 : Find the maximal α such that there is x with $\omega^{\epsilon_0+2} = \phi_\alpha(x)$. We check: $\omega^{\epsilon_0+2} = \phi_0(\epsilon_0 + 2)$. Since also $\epsilon_0 + 2 \neq \phi_0(\epsilon_0 + 2)$, there is no such x for $\alpha = 1$. So we take $\alpha = 0$ and $x = \epsilon_0 + 2$. The expansion so far is $\phi_0(\epsilon_0 + 2)$

Step 3 : Find the maximal y such that $\omega^y \leq \epsilon_0 + 2$, the maximal n such that $\omega^y \cdot n \leq \epsilon_0 + 2$ and the maximal z such that $\omega^y \cdot n + z \leq \epsilon_0 + 2$. Observe that, since $\epsilon_0 = \omega^{\epsilon_0}$, we have $y = \epsilon_0$, $n = 1$ and $z = 2$. The expansion so far is $\phi_0(\omega^{\epsilon_0} + 2)$.

Step 4 : Find the maximal α such that there is x with $\omega^{\epsilon_0} = \phi_\alpha(x)$. We check: $\omega^{\epsilon_0} = \phi_0(\epsilon_0) = \phi_1(0)$. Since also $0 \neq \phi_1(0)$, there is no such x for $\alpha = 2$. So we take $\alpha = 1$ and $x = 0$. The expansion so far is $\phi_0(\phi_1(0) + 2)$.

Step 5 : Find the maximal y such that $\omega^y \leq 2$, the maximal n such that $\omega^y \cdot n \leq 2$ and the maximal z such that $\omega^y \cdot n + z \leq 2$. Trivially, $y = 0$, $n = 2$ and $z = 0$. The expansion so far is $\phi_0(\phi_1(0) + \omega^0 + \omega^0)$

Step 6 : Find the maximal α such that there is x with $\omega^0 = \phi_\alpha(x)$. We check: $\omega^0 = \phi_0(0)$. Since also $0 \neq \phi_0(0)$, there is no such x for $\alpha = 1$. So we take $\alpha = 0$ and $x = 0$ The expansion so far is $\phi_0(\phi_1(0) + \phi_0(0) + \phi_0(0))$

Step 7 : Find the maximal y such that $\omega^y \leq 1$, the maximal n such that $\omega^y \cdot n \leq 1$ and the maximal z such that $\omega^y \cdot n + z \leq 1$. Trivially, $y = 0$, $n = 1$ and $z = 0$. The expansion so far is $\phi_0(\phi_{\omega^0}(0) + \phi_0(0) + \phi_0(0))$

Step 8 : Find the maximal α such that there is x with $\omega^0 = \phi_\alpha(x)$. We check: $\omega^0 = \phi_0(0)$. Since also $0 \neq \phi_0(0)$, there is no such x for $\alpha = 1$. So we take $\alpha = 0$ and $x = 0$. The expansion so far is $\phi_0(\phi_{\phi_0(0)}(0) + \phi_0(0) + \phi_0(0))$. We are done.

References

[1] Jeremy Avigad and Erich H. Reck, "Clarifying the nature of the infinite": the development of metamathematics and proof theory, Carnegie-Mellon Technical Report CMU-PHIL-120, 2001. Also available as http://www.andrew.cmu.edu/user/avigad/Papers/infinite.pdf

[2] N. Bourbaki, Sur le theorème de Zorn, Arch. Math. **2** (1950), 434–437.

[3] Buchholz, W. and Wainer S. Provably computable functions and the fast-growing hierarchy. Logic and Combinatorics, AMS Contemporary Mathematics **65** (1985) pp 179-198.

[4] John Doner and Alfred Tarski. An Extended Arithmetic of ordinal numbers. Fundamenta Mathematicæ **LXV**(1969) 95–127. Also on `http://www.math.ucsb.edu/~doner/articles/`.

[5] Fairtlough, M and Wainer, S. Hierarchies of Provably Recursive Functions: chapter III of Handbook of Proof Theory (S.Buss Ed) Elsevier (1995) pp 148-205.

[6] Torkel Franzen, Inexhaustibility: A Non-Exhaustive Treatment, Lecture Notes in Logic **16**, A. K. Peters, Ltd., 2004.

[7] Solomon Feferman, Highlights in proof theory, in Proof Theory, eds. V. F. Hendricks et al, Kluwer, Dordrecht (2000), pp. 11-31. Also available at http://math.stanford.edu/ feferman/papers.html

[8] Solomon Feferman, Proof theory since 1960, prepared for the Encyclopedia of Philosophy Supplement, Macmillan Publishing Co., New York. Also available at http://math.stanford.edu/ feferman/papers.html

[9] Geza Fodor, Eine Bemerking zur Theorie der regressiven Funktionen, Acta Scientiarum Mathematicarum, Szeged, **17** (1956), pp. 139–142.

[10] Gerhard Gentzen, Die Widerspruchfreiheit der reinen Zahlentheorie, Mathematische Annalen **112** (1936), 493-565. Translated as "The consistency of arithmetic" in M. E. Szabo ed., The Collected Works of Gerhard Gentzen, North-Holland, Amsterdam, 1969.

[11] R. Goodstein, On the restricted ordinal theorem, Journal of Symbolic Logic, **9** (1944), 33-41.

[12] Hardy, G. H. Quarterly J. of Pure and Applied Mathematics. **35** (1903) 87-94.

[13] A Joyal and E Moerdijk Algebraic set theory. LMS lecture notes series **220** CUP 1995

[14] L. Kirby and J. Paris, Accessible independence results for Peano arithmetic, Bull. London. Math. Soc. **14** (1982), 285–93.

[15] W.V. Quine Set Theory and its Logic. Harvard

[16] Diana Schmidt. Built-up systems of fundamental sequences and hierarchies of number-theoretic functions. Arch. Math. Logik. **18** 4753 (???) 1976.

[17] H. Simmons A Comparison of Two Systems of Ordinal Notation, Arch. Math. Logic **43** (2004) pp 65–83.

[18] Alan M. Turing, Systems of logic defined by ordinals, Proc. London Math. Soc., Series 2, **45** (1939), 161–228.

[19] Wainer, S. [1996] Basic Proof Theory with Applications to Computation. Leeds preprint series **18**.

[20] Wikipedia, Ordinal numbers, `http://en.wikipedia.org/wiki/Ordinal_number`
Ordinal arithmetic, `http://en.wikipedia.org/wiki/Ordinal_arithmetic`
Large countable ordinals, `http://en.wikipedia.org/wiki/Large_countable_ordinals`

[21] E. Witt, Beweisstudien zum Satz von M. Zorn, Math. Nachr. **4** (1951), 434–438.

Material which will eventually turn into a proper definition of the ordinals as a higher-order rectype

We define On, $<_{On}$ and $=_{On}$ by a simultaneous recursion

This material on finding the correct declaration of *On* as a recursive datatype is still experimental.

Definition 36 *0 is an ordinal;*
If α *is an ordinal, so is* $\texttt{succ}(\alpha)$*;*
If X *is a set of ordinals, then* $\texttt{sup}(X)$ *is an ordinal;*
$0 \leq \alpha$;
$\alpha < \texttt{succ}(\alpha)$;
$(\forall\alpha \in X)(\exists\beta \in Y)(\alpha \leq \beta) \to \texttt{sup}(X) \leq \texttt{sup}(Y)$;
$\alpha \in X \to \alpha \leq \texttt{sup}(X)$;

and various boring axioms to make trivial facts obvious:

$\alpha \leq \beta \to \beta \leq \alpha \to \alpha = \beta$;
$\alpha \leq \beta \to \beta \leq \gamma \to \alpha \leq \gamma$;
$\alpha < \beta \to \beta < \gamma \to \alpha < \gamma$;
$\alpha = \beta \to \beta = \gamma \to \alpha = \gamma$;
$\alpha < \beta \to \beta < \alpha \to \perp$;
$\alpha = \beta \to \beta = \alpha$;
$\alpha = \beta \to \alpha \leq \beta$;
$\alpha < \beta \to \alpha \leq \beta$.
$(\forall\alpha \in S_1)(\exists\beta \in S_2)(\alpha \leq \beta) \to \texttt{sup}(S_1) \leq \texttt{sup}(S_2)$

(I omit the—even more boring—obvious axioms to the effect that = is a congruence relation for the other relations. Omitted too—for the moment—are axioms to characterise `sup`: I'm thinking of things like $(\forall\alpha)(\forall S \subset On)((\alpha < \texttt{sup}(S)) \to (\exists\beta \in S)(\alpha \leq \beta))$ which i think is Horn.)

In the above definition $\leq$ and $<$ are a pair of partial order/strict partial order. I have exploited both notations in order to ensure that all the clauses in the declaration are Horn and we thereby have a legitimate datatype declaration.

The strict order $<$ (I have omitted the subscript) is the **engendering relation** of the datatype of ordinals. It is wellfounded for the usual reasons.

Once one has equipped On with a wellorder, one can use ideas like that of *order topology*. Conveniently and unsurprisingly it turns out that this gives us the same notion of limit as we presupposed in the extra constructor `sup`. The notion of continuous function will of course be important to us.

H I A T U S

Or do we define $\leq$ and On by a simultaneous recursion and define $=$ as $\leq \geq$?

Appendix 1: The engendering relation on On is a wellorder

THEOREM 37 *$<_{On}$ is a wellorder.*

Proof:

The engendering relation on ordinals is a wellfounded partial order—for the usual reasons; the hard part is showing that it is a *total* order.

The proof was discovered simultaneously and independently by Witt [21] and Weil [2] (tho' neither of these two gentlemen would have described it in those terms[8]) and was used by them to establish that every inflationary function f from a chain-complete poset with a bottom element into itself has a fixed point. The proof proceeds by considering the inductively defined set containing the bottom element, closed under f and suprema of chains. The part of the proof that concerns us here is the proof that this object is a chain. This of course is simply a proof that the ordinals are wellordered by $<$. All I have done is recast their argument as a proof of this fact about ordinals.

Let us say an ordinal α is **normal** if

$$(\forall\beta)(\beta < \alpha \to \mathtt{succ}(\beta) \leq \alpha).$$

If α is normal, then we prove by induction on 'β' that

$$(\forall\beta)(\beta \leq \alpha \vee \mathtt{succ}(\alpha) \leq \beta).$$

That is to say, we show that, if α is normal, then

$$\{\beta : \beta \leq \alpha \vee \mathtt{succ}(\alpha) \leq \beta\}$$

contains 0 and is closed under `succ` and `sup`s of chains and is therefore a superset of On. Let us deal with each of these in turn.

1. (Contains 0); By stipulation.

[8]Thanks to Peter Johnstone for showing me this material.

2. (Closed under $\texttt{succ}$); If $\gamma \in \{\beta : \beta \leq \alpha \vee \texttt{succ}(\alpha) \leq \beta\}$, then either

   (a) $\gamma < \alpha$, in which case $\texttt{succ}(\gamma) \leq \alpha$ by normality of α and $\texttt{succ}(\gamma) \in \{\beta : \beta \leq \alpha \vee \texttt{succ}(\alpha) \leq \beta\}$; or

   (b) $\gamma = \alpha$, in which case $\texttt{succ}(\alpha) \leq \texttt{succ}(\gamma)$ so $\texttt{succ}(\gamma) \in \{\beta : \beta \leq \alpha \vee \texttt{succ}(\alpha) \leq \beta\}$; or

   (c) $\texttt{succ}(\alpha) \leq \gamma$, in which case $\texttt{succ}(\alpha) \leq \texttt{succ}(\gamma)$ (because $\texttt{succ}$ is inflationary) and $\texttt{succ}(\beta) \in \{\beta : \beta \leq \alpha \vee \texttt{succ}(\alpha) \leq \beta\}$.

3. (Closed under $\texttt{sup}$s of chains); Let $S \subseteq \{\beta : \beta \leq \alpha \vee \texttt{succ}(\alpha) \leq \beta\}$ be a chain. If $(\forall \beta \in S)(\beta \leq \alpha)$, then $\texttt{sup}(S) \leq \alpha$. On the other hand, if there is $\beta \in S$ s.t. $\beta \not\leq \alpha$, we have $\texttt{succ}(\alpha) \leq \beta$ (by normality of α); so $\texttt{sup}(S) \geq \texttt{succ}(\alpha)$ and $\texttt{sup}(S) \in \{\beta : \beta \leq \alpha \vee \texttt{succ}(\alpha) \leq \beta\}$.

Next we show that everything in On is normal. Naturally we do this by induction: the set of normal ordinals will contain 0 and be closed under $\texttt{succ}$ and $\texttt{sup}$s of chains.

1. (Contains 0); 0 is clearly normal.

2. (Closed under $\texttt{succ}$); Suppose $\alpha \in \{\gamma : (\forall \beta)(\beta < \gamma \rightarrow \texttt{succ}(\beta) \leq \gamma\}$. We will show $(\forall \beta)(\beta < \texttt{succ}(\alpha) \rightarrow \texttt{succ}(\beta) \leq \texttt{succ}(\alpha))$. So assume $\beta < \texttt{succ}(\alpha)$. This gives $\beta \leq \alpha$ by normality of α. If $\beta = \alpha$, we certainly have $\texttt{succ}(\beta) \leq \texttt{succ}(\alpha)$, as desired, and if $\beta < \alpha$, we have $\texttt{succ}(\beta) \leq \alpha \leq \texttt{succ}(\alpha)$.

3. (Closed under $\texttt{sup}$s of chains); Suppose $S \subseteq \{\gamma : (\forall \beta \in On)(\beta < \gamma \rightarrow \texttt{succ}(\beta) \leq \gamma)\}$ is a chain. If $\beta < \texttt{sup}(S)$, we cannot have $(\forall \gamma \in S)(\beta \geq \texttt{succ}(\gamma))$ for otherwise $(\forall \gamma \in S)(\beta \geq \gamma)$ (by transitivity of $<$ and inflationarity of $\texttt{succ}$), so for at least one $\gamma \in S$ we have $\beta \leq \gamma$. If $\beta < \gamma$, we have $\texttt{succ}(\beta) \leq \gamma \leq \texttt{sup}(S)$ since γ is normal. If $\beta = \gamma$, then γ is not the greatest element of S, so in S there is $\gamma' > \gamma$ and then $\texttt{succ}(\beta) \leq \gamma' \leq \texttt{sup}(S)$ by normality of γ'.

If α and β are two things in On, we have $\beta \leq \alpha \vee \texttt{succ}(\alpha) \leq \beta$ by normality of α, so the second disjunct implies $\alpha \leq \beta$, whence $\beta \leq \alpha \vee \alpha \leq \beta$. So On is a chain as promised,

5 Answers to selected exercises

Exercise 4

PTJ comments:

[*The original question worked with ω-continuous functions, for which one has a much easier proof of the existence of fixed points, but the question itself becomes harder because you have to verify that every function in sight is ω-continuous. As it stands, it should be pretty easy, except for the proof that m is order-preserving* *(needed to show that f is order-preserving): for this, observe that if $x_1 \le x_2$ then $m(x_2)$ is a 'post-fixed point' of g_{x_1} (that is, $m(x_2) \ge g_{x_1}(m(x_2))$), and so $\{y \in Q : y \le m(x_2)\}$ is a 'closed set' in the sense used in the construction of the least fixed point $m(x_1)$ of g_{x_1}.*]

And my discussion. . .

1. Given $x \in P$, suppose $y_1 \le y_2 \in Q$. Then $\langle x, y_1\rangle \le \langle x, y_2\rangle$ so $h(\langle x, y_1\rangle) \le h(\langle x, y_2\rangle)$ so $g_x(y_1) = h_2(\langle x, y_1\rangle) \le h_2(\langle x, y_2\rangle) = g_x(y_2)$, so g_x is order-preserving.

2. m is order-preserving. *Proof:* Suppose $x_1 \le x_2 \in P$. Set $Y = \{y \in Q : y \le m(x_2)\}$. For $y \in Y$ we have $g_{x_1}(y) = h_2(\langle x_1, y\rangle) \le h_2(\langle x_2, y\rangle) = m(x_2)$ so g_{x_1} acts on Y.

   Now if $C \subseteq Y$ is a chain then its sup is below $m(x_2)$ so Y is chain-complete, whence $g_{x_1}\restriction Y$ has a fixed point $y_0 \in Y$ with $m(x_1) \le y_0 \le m(x_2)$ and m is order-preserving.

   Now suppose $x_1 \le x_2 \in P$ again. Then $\langle x_1, m(x_1)\rangle \le \langle x_1, m(x_1)\rangle$. So $f(x_1) \le f(x_2)$ and f is order-preserving.

3. $h(x_0, m(x_0)) = \langle f(x_0), g_{x_0}(m(x_0))\rangle = \langle x_0, m(x_0)\rangle$ so $\langle x_0, m(x_0)\rangle$ is a fixed point of h. Let $\langle x, y\rangle$ be the least fixed point of h. Then $g_x(y) = y$ so $y \ge m(x)$.

4. Let $Z = \{\langle a, b\rangle \in P \times Q : \langle a, b\rangle \le \langle x, m(x))\rangle\}$

   For $\langle a, b\rangle \in X$:

   $h(a, b) \le h(x, m(x)) \le \langle h_1(x, y), h_2(x, m(x))\rangle = \langle x, m(x)\rangle$ so h acts on Z. Furthermore, Z is chain-complete and has a least element, similar to claim above. So h has a fixed point $\langle x', y'\rangle \in Z$. Now $\langle x, y\rangle \le \langle x', y'\rangle \le \langle x, m(x)\rangle$. But $\langle x, y\rangle \ge \langle x, m(x)\rangle$ so $\langle x, y\rangle = \langle x, m(x)\rangle$.

   So x is a fixed point for f, whence $x \ge x_0$. But m is order-preserving so $y = m(x) \ge m(x_0)$. So $\langle x, y\rangle \ge \langle x_0, m(x_0)\rangle$ and $\langle x_0, m(x_0)\rangle$ is the least fixed point of h.