

Logic and Set Theory:
Dr Russell's Example Sheets for 2019

Thomas Forster

May 10, 2020

Chapter 1

Dr Russell's Example Sheet for 2019

That's *Doctor* Russell, not *Earl* Russell.

Our Promise To You...
This Stuff Will Do Your Head In

Or your money back!

A few of these answers come from Qiaochu Yuan, from years ago, at an earlier appearance of (some of) these questions. (Some of you may have heard of him: he has turned into a *wunderkind* with an internet presence.) He is not responsible for their present form, co's i've hacked them about.

...and not terribly well, co's Dr Russell has tweaked some of the legacy questions and i haven't tweaked the answers properly—yet.

Sheet 1

Question 1

Which of the following propositional formulæ are tautologies?

- (i) $((p \rightarrow (q \rightarrow r)) \rightarrow (q \rightarrow (p \rightarrow r)))$;
- (ii) $((p \rightarrow q) \rightarrow r) \rightarrow ((q \rightarrow p) \rightarrow r)$;
- (iii) $((p \rightarrow q) \rightarrow p) \rightarrow p$;
- (iv) $(p \rightarrow (p \rightarrow q)) \rightarrow p$.

(i)

Suppose that the proposition evaluates to 0 under some valuation ν . Then $\nu(p_1 \rightarrow (p_2 \rightarrow p_3)) = 1$ and $\nu(p_2 \rightarrow (p_1 \rightarrow p_3)) = 0$, whence $\nu(p_2) = 1, \nu(p_1 \rightarrow p_3) = 0$, whence $\nu(p_1) = 1, \nu(p_3) = 0$. It follows that $\nu(p_2 \rightarrow p_3) = 0$, whence finally $\nu(p_1 \rightarrow (p_2 \rightarrow p_3)) = 0$; contradiction. So the proposition is a tautology.

(ii)

Let $\nu(p_1) = 1, \nu(p_2) = \nu(p_3) = 0$. Then $\nu(p_2 \vee p_3) = 0, \nu(p_1 \vee p_2) = 1, \nu(p_1 \vee p_3) = 1$, whence $\nu((p_1 \vee p_2) \wedge (p_1 \vee p_3)) = 1$ and

$$\nu(((p_1 \vee p_2) \wedge (p_1 \vee p_3)) \rightarrow (p_2 \vee p_3)) = 0.$$

(iii)

Suppose that the proposition evaluates to 0 under some valuation ν . Then $\nu(p_1 \rightarrow (\neg p_2)) = 1$ and $\nu(p_2 \rightarrow (\neg p_1)) = 0$, whence $\nu(p_2) = 1, \nu(\neg p_1) = 0, \nu(p_1) = 1$. But this implies $\nu(p_1 \rightarrow (\neg p_2)) = 0$; contradiction. So the proposition is a tautology.

Question 2

[PTJ sez (inter alia) *The fact that $\{\neg p\} \vdash (p \rightarrow q)$ is needed in the proof of the Completeness Theorem.*]

QY supplies this proof.

By the deduction theorem, it suffices to show that $\perp \vdash q$. The following is a proof:

t_1	\perp	(in S)
t_2	$\perp \rightarrow ((q \rightarrow \perp) \rightarrow \perp)$	K
t_3	$(q \rightarrow \perp) \rightarrow \perp$	(modus ponens from t_1, t_2)
t_4	$((q \rightarrow \perp) \rightarrow \perp) \rightarrow q$	(axiom 3)
t_5	q	(modus ponens from t_3, t_4)

Then by the proof of the deduction theorem, the following is a proof that $\perp \rightarrow q$:

1	$\perp \rightarrow (\perp \rightarrow \perp)$	K
2	$\perp \rightarrow ((\perp \rightarrow \perp) \rightarrow \perp)$	K
3	$(\perp \rightarrow ((\perp \rightarrow \perp) \rightarrow \perp)) \rightarrow ((\perp \rightarrow (\perp \rightarrow \perp)) \rightarrow (\perp \rightarrow \perp))$	S
4	$(\perp \rightarrow (\perp \rightarrow \perp)) \rightarrow (\perp \rightarrow \perp)$	(modus ponens from 2, 3)
5	$\perp \rightarrow t_1$	(modus ponens from 1, 4)
6	t_2	K
7	$t_2 \rightarrow (\perp \rightarrow t_2)$	K
8	$\perp \rightarrow t_2$	(modus ponens from 6, 7)
9	$(\perp \rightarrow t_2) \rightarrow ((\perp \rightarrow t_1) \rightarrow (\perp \rightarrow t_3))$	S
10	$(\perp \rightarrow t_1) \rightarrow (\perp \rightarrow t_3)$	(modus ponens from 8, 9)
11	$\perp \rightarrow t_3$	(modus ponens from 5, 10)
12	t_4	(axiom 3)
13	$t_4 \rightarrow (\perp \rightarrow t_4)$	K
14	$\perp \rightarrow t_4$	(modus ponens from 12, 13)
15	$(\perp \rightarrow t_4) \rightarrow ((\perp \rightarrow t_3) \rightarrow (\perp \rightarrow t_5))$	S
16	$(\perp \rightarrow t_3) \rightarrow (\perp \rightarrow t_5)$	(modus ponens from 14, 15)
17	$\perp \rightarrow t_5$	(modus ponens from 11, 16).

Question 3

We want to show that $p \vdash (p \rightarrow \perp) \rightarrow \perp$. By the deduction theorem, it suffices to show that $\{p, p \rightarrow \perp\} \vdash \perp$. But this follows by *modus ponens*.

Question 4

We want to show that $\{p, q\} \vdash (p \rightarrow (q \rightarrow \perp)) \rightarrow \perp$.

(i) By the deduction theorem, it suffices to show that $\{p, q, p \rightarrow (q \rightarrow \perp)\} \vdash \perp$. But this follows by two applications of *modus ponens*.

(ii) By the completeness theorem, it suffices to consider a valuation ν with $\nu(p) = \nu(q) = 1$. Then $\nu(q \rightarrow \perp) = 0$, whence $\nu(p \rightarrow (q \rightarrow \perp)) = 0$, from which it follows that $\nu((p \rightarrow (q \rightarrow \perp)) \rightarrow \perp) = 1$.

(iii) By the proof of the deduction theorem, the following is a proof that $\{p, q\} \vdash p \wedge q$, where $x = (p \rightarrow (q \rightarrow \perp))$:

- | | |
|---|-----------------------------|
| (1) $x \rightarrow (x \rightarrow x)$ | K |
| (2) $x \rightarrow ((x \rightarrow x) \rightarrow x)$ | K |
| (3) $(x \rightarrow ((x \rightarrow x) \rightarrow x)) \rightarrow ((x \rightarrow (x \rightarrow x)) \rightarrow (x \rightarrow x))$ | S |
| (4) $(x \rightarrow (x \rightarrow x)) \rightarrow (x \rightarrow x)$ | (modus ponens from 2, 3) |
| (5) $x \rightarrow x$ | (modus ponens from 1, 4) |
| (6) p | (in S) |
| (7) $p \rightarrow (x \rightarrow p)$ | K |
| (8) $x \rightarrow p$ | (modus ponens from 6, 7) |
| (9) q | (in S) |
| (10) $q \rightarrow (x \rightarrow q)$ | K |
| (11) $x \rightarrow q$ | (modus ponens from 9, 10) |
| (12) $(x \rightarrow x) \rightarrow ((x \rightarrow p) \rightarrow (x \rightarrow (q \rightarrow \perp)))$ | S |
| (13) $(x \rightarrow p) \rightarrow (x \rightarrow (q \rightarrow \perp))$ | (modus ponens from 5, 12) |
| (14) $x \rightarrow (q \rightarrow \perp)$ | (modus ponens from 8, 13) |
| (15) $(x \rightarrow (q \rightarrow \perp)) \rightarrow ((x \rightarrow q) \rightarrow (x \rightarrow \perp))$ | S |
| (16) $(x \rightarrow q) \rightarrow (x \rightarrow \perp)$ | (modus ponens from 14, 15) |
| (17) $x \rightarrow \perp$ | (modus ponens from 11, 16). |

Now, from the premise $\neg p$, (or $p \rightarrow \perp$), together with a proof that $\perp \rightarrow q$ for arbitrary q , we conclude that $p \rightarrow q$ by the example in class.

(Qiaochu Yuan again)

Question 5

It suffices to set $q := \neg p$. Suppose there were a valuation ν such that $\nu((p \rightarrow \neg p) \rightarrow \neg(\neg p \rightarrow p)) = 0$. Then $\nu(p \rightarrow \neg p) = 1$ and $\nu(\neg(\neg p \rightarrow p)) = 0$, whence $\nu(\neg p \rightarrow p) = 1$. But if $\nu(p) = 1$, then the first condition is impossible, and if $\nu(p) = 0$, then the second condition is impossible; contradiction. So there exists no such valuation.

Question 6

Let P, Q, R be three consistent and deductively closed sets—the beliefs of the three parties. Then it is not possible to prove \perp from any of P, Q, R , whence it follows that it is not possible to prove \perp from any subset of any of P, Q, R ; in particular it is not possible to prove \perp from $P \cap Q \cap R$. It follows that $P \cap Q \cap R$ is consistent. Similarly, if t is a proposition which can be proven from $P \cap Q \cap R$, then it can be proven from P or Q or R , so it is in $P \cap Q \cap R$. It follows that $P \cap Q \cap R$ is deductively closed.

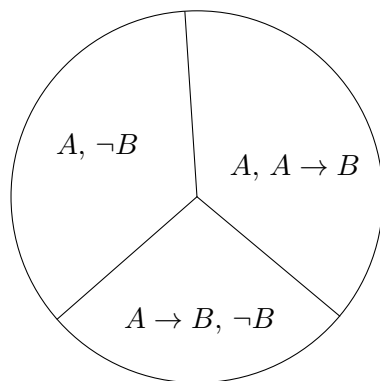
However, if P, Q and R are three consistent deductively closed sets of propositions, there is no guarantee that $(P \cap Q) \cup (P \cap R) \cup (Q \cap R)$ is deductively closed or consistent. For consider:

P is the deductive closure of $\{A, \neg B\}$

Q is the deductive closure of $\{A, A \rightarrow B\}$

R is the deductive closure of $\{A \rightarrow B, \neg B\}$

A majority now believe $A, A \rightarrow B, \neg B$. This is not consistent. And, since the majority doesn't believe \perp , it isn't deductively closed either.



Observe (this is a check on your comprehension) that this can be extended to any finite number of sets—asking for larger majorities doesn't change anything. Divide the world into four bundles. Bundles 1, 2 and 3 all believe A ; bundles 2, 3, 4 all believe $A \rightarrow B$; bundles 3, 4 and 1 all believe $B \rightarrow C$; finally bundles 4, 1 and 2 all believe $\neg C$. Each bundle has consistent beliefs but the beliefs held by a 3/4 majority are not consistent.

Mind you, if you have *infinitely* many people the then set of things believed by *cofinitely many* of them is consistent!

Question 7

We can prove by induction that if A is derivable from K and S and contains \perp then \perp can be replaced in A —and indeed throughout the proof of A —by some new letter not in the proof of A ; the transformed proof is still a proof within the meaning of the act. So the modified A is still deducible from K and S . However the result of modifying the third axiom in this way is not a propositional tautology, and therefore cannot be deduced from K and S .

Actually one—no, *three*—of my supervisees came up with this. I had neglected to tell them it wouldn't work, so they just went ahead and did it, and it worked. Quite embarrassing really.

They say: Read ' $p \rightarrow q$ ' as ' $(\neg p) \wedge q$ ', read ' \perp ' as ' \perp ' and take the designated truth-value to be 0. Then axioms K and S (aka 1 and 2) always take truth-value 0, and MP preserves the property of always taking truth-value 0. That ensures that axiom 3 does not take the designated value.

Question 8

Suppose not ...

Consider $\{\neg t_n : n \in \mathbb{N}\}$. This is an inconsistent theory, since every v makes at least one t_n true. So by compactness there is a N such that $\{\neg t_n : n < N\} \vdash \perp$. But that is to say that every valuation must make true one of the t_n with $n < N$.

Why is the compactness theorem for propositional logic like the compactness of the space of valuations? The space of valuations is compact. That is beco's it is the product of lots of copies of the two-point space (one copy for each propositional letter) and the two-point space is compact. And a product of compact spaces is compact. (That's Tikhonov—in fact a subtly weaker version of Tikhonov that sez that a product of compact *Hausdorff* spaces is compact *Hausdorff*). For any propositional formula ϕ the set $[[\phi]]$ of valuations making it true is closed (in fact clopen). Suppose now that Γ is an inconsistent set of formulæ. Then $\{[[\phi]] : \phi \in \Gamma\}$ is a family of closed sets with empty intersection. So some finite subset of it has empty intersection. So there is a finite $\Gamma' \subseteq \Gamma$ with $\Gamma' \models \perp$.

Question 9

The thought here is that a transfinite *monotone* process—even if it is nondeterministic—might complete. (You need monotonicity, since without it you have no handle on what a limit stage looks like. ... look at q 13 on <https://www.dpmms.cam.ac.uk/study/II/Logic/2018-2019/logic20192.pdf> to get a flavour of quite how horrible things can be if you don't have monotonicity). You need the axiom of choice if it is nondeterministic, but Hartogs' lemma ensures that you never run out of stages: if the process never completes it's not beco's you have run out of stages but beco's the process is internally inconsistent.

Question 10

This is a lovely result: *The Interpolation Lemma*. I wrote up a proof in www.dpmms.cam.ac.uk/~tf/chchlectures.pdf (it's a searchable .pdf) and i don't want to repeat myself.

Mind you, i could supply this argument (which is *not* in chchlectures.pdf) and which one of you came up with. Suppose $a \models c$, and let p be a propositional letter that occurs in a but not in c .

Then we have

$$a$$

iff

$$a \wedge (p \vee \neg p)$$

iff

$$(a \wedge p) \vee (a \wedge \neg p)$$

We can simplify $a \wedge p$ by replacing every occurrence of ‘ p ’ in a by \top , and we can simplify $a \wedge \neg p$ by replacing every occurrence of ‘ p ’ in a by \perp . But now we have a consequence of a that entails c —just as a did—but has one fewer letter alien to c .

I do want to make the point tho’, that that nice proof does not generalise to the first-order case. In the first-order case one has only the compactness argument. Let X be the set of consequences of a that can be expressed in the common vocabulary. Persuade yourself that c follows from X , and therefore (by compactness) from some finite subset $X' \subseteq X$. Then the desired b is $\bigwedge X'$.

This is actually an old tripos question: 2013:p3:18G, and probably earlier too.

Question 11

Any finite set of sentences has an independent subset. You can discard a sentence that follows from the remaining sentences. You can do this deterministically or non-deterministically, it doesn’t matter. It doesn’t matter in the sense that you will get an independent subset whatever happens, but which independent subset you get might depend on the order in which you do your weeding. For example if you start with $\{p, p \longleftrightarrow q, q\}$ you can drop any one of the three to obtain an independent subset. (This is a repurposing of a standard illustration of three events any two of which are independent; you may know it from elsewhere).

Let the propositional alphabet P be $\{p_i : i \in \mathbb{N}\}$.

Then the set $\{\bigwedge_{i \leq n} p_i : n \in \mathbb{N}\}$ is a set of formulæ with no equivalent independent subset.

For the second part, suppose $\{A_i : i \in \mathbb{N}\}$ axiomatises a theory T . Perform a *weeding* operation by removing any A_i that follows from $\{A_j : j < i\}$. Then renumber.

Next consider the axioms

$$B_i := \left(\bigwedge_{j < i} A_j\right) \rightarrow A_i.$$

(Observe that B_1 is just A_1 —beco’s the empty conjunction is just the **true**). Clearly the B_i axiomatise T . We will show that they are independent.

Fix i and consider B_i , which is $(\bigwedge_{j < i} A_j) \rightarrow A_i$. Beco’s of the weeding it is not a tautology. So there is a valuation making it false. Any such valuation both

- (i) makes A_j true for $j < i$ (and thereby makes all the B_j with $j < i$ true by making the consequents true) and
- (ii) makes A_i false (and thereby makes true all the B_k with $k > i$ by making all their antecedents false).

Thus, for every i , there is a valuation making B_i false and all the other B_j true. So no B follows from any of the others.

Question 12

The answer is ‘no’ but there is no obvious reason to expect it. If you wanted to guess that the answer is ‘no’ you could reflect that the collection of deductive consequences of the first two axioms using *modus ponens* is an inductively defined set and so supports a kind of induction, so you might try to find some property possessed by the first two axioms that is preserved by *modus ponens* that is not possessed by some special tautology. And this is in fact exactly what we will do.

The counterexample is $((A \rightarrow B) \rightarrow A) \rightarrow A$, commonly known as *Peirce’s law*. Easy to check that it is a tautology...less easy to see that it does not follow from K and S .

Axiom K : $A \rightarrow (B \rightarrow A)$.

Axiom S : $(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$.

The idea that is key to cracking this question is the thought that there might be more than one notion of validity, *i.e.*, there might be some other property that is possessed by K and S and which is preserved by *modus ponens* but is not possessed by some (unspecified and so far undiscovered) tautology containing only ‘ \rightarrow ’. There is a ready supply of these notions in the form of *many-valued truth-tables*. We will use the following three-valued truth-table for the connective ‘ \rightarrow ’.

\rightarrow	1	2	3
1	1	2	3
2	1	1	3
3	1	1	1

For our purposes, think of truth-value 1 as **true** and the other two truth-values as two flavours of **false**.

Notice that, in this truth table, if A and $A \rightarrow B$ both take truth-value 1, so does B . Notice also that K and S take truth-value 1 under all assignments of truth-values to the letters within them. So if ϕ is deducible from K and S , it must take value 1 under any assignment of truth-values to the literals within it (by structural induction on the family of proofs).

Then check that, if A is given truth-value 2 and B is given truth-value 3, $((A \rightarrow B) \rightarrow A) \rightarrow A$ then gets truth-value 2, rather than 1.

So Peirce’s law is not deducible from K and S .

(Notice that if we ignore the truth-value 2 (so that we discard the second row and the second column) what remains is a copy of the ordinary two-valued table, with 3 as **false** and 1 as **true**. Also, if we similarly ignore the truth-value 3 what remains is a copy of the ordinary two-valued table with 1 as **true** and 2 as **false**.)

This three-valued logic caper looks entirely *ad hoc*—and indeed it is. Or was. Originally. It turned out later that the funny truth-values have genuine mathematical meaning. (Something to do with possible world semantics). But that wasn’t clear to the people who dreamt them up. There’s a moral there ... (If you want to know about possible world semantics look at the chapter in www.dpmms.cam.ac.uk/~tf/chchlectures.pdf)

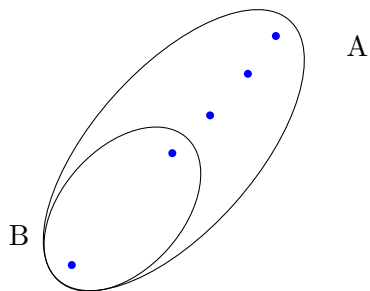
The other moral of this example is that some kinds of mathematics really need formalisation. Unless we had a concept of proof, and of proof by induction on the structures of proofs, we would have no way of demonstrating that $((A \rightarrow B) \rightarrow A) \rightarrow A$ cannot be derived from K and S .

There is a more subtle, more beautiful and more enlightening—but much harder—proof using Curry-Howard, but we probably won’t get round to it. However, if we *did* get round to talking

about Curry-Howard in the supervision then the remainder of this section will make sense to you. I wrote it up from a brief paragraph in an article of Scott's¹ partly for my own good, and it may well benefit from critical eyes such as yours, Dear Reader.

Dana Scott's clever proof

Suppose *per impossibile* that there were a uniformly definable (and, accordingly, invariant) function P for Peirce's law. Let B be a two-membered set, and let A be obtained from B by adding three new elements.



A has five members and B has two, so any function $A \rightarrow B$ identifies a distinguished member of B , namely the one with larger preimage. This defines a function from $A \rightarrow B$ to B , which is to say (since $B \subseteq A$) a function from $A \rightarrow B$ to A . So what we have, in this rather special case, is a distinguished function $(A \rightarrow B) \rightarrow A$. Let us call this function F . F exists only because of the special circumstances we have here contrived, and it's not the sort of thing that P would normally expect to have to deal with, so we should expect P to experience difficulty with it ... which of course is exactly what we want! But, if we have a term P , we can apply it to F to obtain a distinguished member of A . But clearly there is no way of picking a member of A in this way. The alleged existence of a uniformly definable P is trying to tell us that whenever we have a set of five things divided into two parts, one with two things in it and the other with three, then one of the five things is distinguished. And that's clearly not true.

On what features of A and B does this counterexample rely? A function $A \rightarrow B$ has to give us (via the pigeonhole principle) a distinguished element of B , so we need B to have two elements, and A (and therefore $A \setminus B$) to have an odd number. $|A \setminus B| = 1$ is no good, beco's then A has a distinguished element, which we don't want. $|A \setminus B| = 3$ is the smallest number that will do, and that is what Dana Scott gives us.

If you want to know more about Curry-Howard (and you should) there are a million and one treatments of it out there. You could start with the chapter in www.dpmms.cam.ac.uk/~tf/chchlectures.pdf.

¹Dana Scott D.S. Semantical Archæology, a parable. In: Harman and Davidson eds, Semantics of Natural Languages. Reidel 1972 pp 666–674.

Sheet 2

Question 1

An old chestnut. You need to take seriously the possibility that $x = y$ and the trouble this might cause for you. It certainly *seems* at first blush that you have to do a case split on whether or not the two components are the same. Remarkably this is not so: this pairing/unpairing kit is constructive!

$$\mathbf{pair}(x, y) = \{\{x\}, \{x, y\}\}.$$

$$\mathbf{fst}(p) = \bigcap \bigcap p \text{ and}$$

$$\mathbf{snd}(p) = \text{the unique member of } \bigcup p \text{ belonging to exactly one member of } p:$$

$$x = \mathbf{snd}(p) \longleftrightarrow (\exists! z)(z \in p \wedge x \in z).$$

(Thanks to Peter Johnstone and Randall Holmes for showing me that **snd** is actually constructive. In fact no negation signs anywhere in sight.)

Ordinals

Most of these following questions on ordinals are routine. There are some things worth bearing in mind. Ordinals are merely a special kind of linear order type, and that is a good context for thinking of them. The associativity of $+$ and \times are facts about order types in general, and you don't need induction to do prove them. The failure of commutativity of \times isn't really a fact about ordinals but arises from the inherent asymmetry of the operation: $\alpha \cdot \beta$ is β copies of α , and the thing being copied and the thing indexing the copies have different rôles, so you shouldn't expect to be able to swap them. The fact that you can in fact swap them in the finite case is one of the many oddities of finiteness. In fact (this has only just occurred to me) you can probably use it to give a definition of finiteness.

A word is in order on the definition of \leq on ordinals and linear order types generally. Dr Russell takes $\alpha \leq \beta$ to mean that a thing of length β is isomorphic to an initial segment of a thing of length α . However there is also a definition that says that $\beta \leq \alpha$ if a thing of length β *can be injected into a thing of length α in an order preserving way*. Altho' these conditions are clearly not equivalent for arbitrary linear order types, they are equivalent for ordinals. Indeed: a total order is a wellorder iff every subordering is iso to an initial segment. You might like to try to prove this fact: it has been an example sheet question at various times in the past.

If (*pace* Dr Russell) you read \leq in the "order-preserving-injection" sense, then things like q 3 can be proved for *all* linear order types *outright*, without induction, just synthetically. You then get a proof of what Dr Russell asked you to prove (for ordinals) by appealing to the fact that the two definitions agree *on ordinals*.

Another reason for avoiding induction wherever possible is that the bookkeeping can be horrendous. Suppose you wanted to prove $\alpha(\beta\gamma) = (\alpha\beta)\gamma$ by induction; how do you do it? You fix α and β and do it by induction on γ ? You prove by induction on α that it's true for all β and γ ? There are probably 27 different ways trying it. Don't go there.

Question 2

Question 3

Question 4

My colleague Peter Smith asked me years ago about the “synthetic” definition of ordinal exponentiation. The significance is that it seems that there cannot be a synthetic definition of the “next” operation after exponentiation. I started writing out an answer, but got bogged down, so i am very glad to be given this prompt.

Let $\langle A, <_A \rangle$ and $\langle B, <_B \rangle$ be wellorderings of length α and β respectively, with their bottom elements notated ‘ 0_A ’ and ‘ 0_B ’ respectively.

A function $f : A \rightarrow B$ is said to be “of finite support” iff it sends all but finitely many of its arguments to 0_B .

The idea is that if α is the order type of $\langle A, <_A \rangle$ and β is the order type of $\langle B, <_B \rangle$ then α^β is the order type of the set of functions $B \rightarrow A$ of finite support, ordered “colex”—by last difference. And let us—strictly in this environment only—write ‘ $B \rightarrow A$ ’ to mean nonstandardly the set of functions from B to A of finite support (rather than—as is customary—the set of *all* functions $B \rightarrow A$).

[Why do we not take a “function of finite support” to be not a total function that sends all but finitely many of its args to 0 but rather one that is undefined except at finitely many arguments? I am not sure, but it’s perhaps beco’s it (slightly) complicates the definition of the colex ordering. Also there is the consideration that exponentiation of this kind is defined also between arbitrary linear order types with a bottom element, not merely between ordinals, and it outputs linear order types. (Notice that to output linear order types it has to restrict itself to functions of finite support; this means that the connection to cardinals is lost). And it looks wrong that this definition of exponentiation for linear order types should work only if the order type has a bottom element. Actually this looks like a good reason for using the partial function definition!]

So let’s write out a proof that the order type of $B \rightarrow A$ [finite support version] really is α^β (where the ordinal exponentiation is defined by the usual recursion). Obviously we fix α (so we are doing a UG at top level) and then run a recursion on the exponent.

My students spent a lot of time proving that the colex ordering on $B \rightarrow A$ is in fact a total ordering. This involves a nasty hacky case analysis, which i think can be sidestepped by fixing A and α , and then proving—by induction on β —that any $B \rightarrow A$ where $\text{otype}(B) = \beta$ is of otype α^β ... which is what i now propose to do.

Let’s take A to be I_α and B to be I_β . (The set of ordinals $< \alpha$ and $< \beta$ respectively).

Base case ... α^0 and α^1 can be done by hand, as it were. Let’s start with α^2 .

Case $\beta = 2$

This matches the usual definition of $\alpha \cdot \alpha$.

Case $\beta = \gamma + 1$

Let us write ‘ C ’ for I_γ ’ to make things readable.

To get α^β consider a γ -shaped skeleton list, waiting to have its locations filled in by ordinals below α , all but finitely many of them 0. Order the results colex, by last difference. Now consider the effect of adding an additional address on the end, so there are now $\gamma + 1$ locations, no longer merely γ . We get lots of new functions $B \rightarrow A$ of finite support. For each $\zeta < \alpha$ we will get a copy of all the old sequences that made up α^β . The old functions from C to A are not total functions B

to A so we have to turn them into such functions if we are to make them into members of $B \rightarrow A$. Best to do that by deeming them to all send γ to 0. We now find that they are duplicated by new functions so we can simply ignore them. However we are interested in their order type, since that is the ghost that remains. Think about the difference between the legacy functions from C , and the new functions that send γ to something other than 0. Since we are ordering things by *last* difference this collection is divided into bundles according to the last element, and each bundle is a copy of the bundle of legacy functions. Each such bundle is of order type α^γ and there are α of them, so the new collection is of order type $\alpha^{\gamma+1} = \alpha^\beta$ as desired.

Case: β is limit.

The key observation here is that every function in $\alpha^{\beta+1}$ appears on the end of everything in α^β so we are talking **end-extensions**, which makes everything continuous. The point is that it is standard that whenever $\langle \mathcal{W}_i : i \in I \rangle$ is a family of wellorderings linearly ordered by end-extension then the order type of the union is the sup of the order types of the \mathcal{W}_i .

It follows from the foregoing that

$$\alpha^{\beta_1+\beta_2} = \alpha^{\beta_1} \cdot \alpha^{\beta_2},$$

but it might be enlightening to have a independent hand-crafted artisan proof such as you might pick up in Camden market. Let's have a go.

We have two wellorders B_1 and B_2 , with B_2 concatenated on the end of B_1 , and a wellorder A . A function f [of finite support] from $B_1 \sqcup B_2$ to A can naturally be thought of as a pair of functions $f_1 : B_1 \rightarrow A$ and $f_2 : B_2 \rightarrow A$. We have to verify that the lexicographic order on $(B_1 \rightarrow A) \times (B_2 \rightarrow A)$ is the same as the colex order on $(B_1 \sqcup B_2) \rightarrow A$. But that is obvious (isn't it...?)

Question 5

Question 6

Question 7

Question 8

A brief comment.... It's easy to show, if α and β are any-old linear order types and you are using Dr. Russell's definition, then there is such a γ . However if you are to prove that it is *unique* you need α to be an ordinal.

Question 9

(Tripos II 93206). For each countable ordinal α , show that there is a subset of \mathbb{R} which is well-ordered (in the usual ordering) and has order-type α . Is there a well-ordered subset of \mathbb{R} (again, in the usual ordering) of order-type ω_1 ?

It works not just for countable ordinals, but any countable order type whatever!

Take any total order of \mathbb{N} . We will define an injection into \mathbb{Q} by recursion on the naturals. Send each natural number as it pops up to, well, the first positive integer if it is to the *right* of stuff already allocated, or the first negative integer if it is to the *left* of stuff already allocated. If it

is between two things already allocated send it to the arithmetic mean of the things its immediate upper and lower neighbours were sent to.

That's the correct way to do it. There is a wrong way to do it, which most people pounce on, and that is to try to do it by induction on countable ordinals. Clearly if i can embed I_α (the ordinals below α) into \mathbb{Q} then i can embed it into each of $\mathbb{Q} \cap (n, n+1)$, and thus i can embed $I_{\alpha \cdot \omega}$ into \mathbb{Q} . However this ceases to be of any use when I reach an ordinal λ s.t. I_λ is closed under multiplication by ω . When you reach such an ordinal you need a *fundamental sequence* for it, a sequence $\langle \lambda_n : n \in \mathbb{N} \rangle$ whose sup is λ . Then you embed I_{λ_i} into $\mathbb{Q} \cap (i, i+1)$. However you have to use the axiom of choice to pick fundamental sequences for all limit ordinals. I shall spare you the details, since you may well have worked them out for yourself. However i will spell out an amusing detail.

Let us suppose that we have—by the above ruse, using AC—obtained a family $\langle f_\alpha : \alpha < \omega_1 \rangle$ where f_α injects I_α into \mathbb{R} in an order-preserving way. Fix a countable ordinal ζ and consider the ω_1 -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!

REMARK 1. *For each countable ordinal γ , the sequence $\langle f_\gamma(\zeta) : \gamma > \zeta \rangle$ is not monotone nonincreasing.*

Proof: Suppose this is not so, so that

$$(\forall \gamma < \gamma' < \omega_1)(\forall \zeta < \omega_1)(f_\gamma(\zeta) \geq f_{\gamma'}(\zeta)). \quad (1.1)$$

Then, for each $\zeta < \omega_1$, the sequence $\langle f_\gamma(\zeta) : \gamma > \zeta \rangle$ of values given to ζ is nonincreasing, and must be eventually constant. This is because *any* non-increasing sequence from I_{ω_1} into \mathbb{R} must be eventually constant. For suppose it is not eventually constant; think about how many different values it must take. For any countable ordinal there would have to be a later countable ordinal that gets sent to a smaller real. That is to say, the set of ordinals at which an actual decrement occurs is unbounded in I_{ω_1} , and the axiom of choice tells us that any such sequence is of length ω_1 . However (as we will show in our answer to the second part of this question) there can be no such sequence of reals.

So there is an eventually constant value given to ζ , which we shall write ' $f_\infty(\zeta)$ '. But now we have $\alpha < \beta \rightarrow f_\infty(\alpha) < f_\infty(\beta)$. (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.² ■

However, that appealed to the second part of the question, which i had better now prove.

For the second part (“can you do the same for ω_1 ?”) ...

There can be no subset of \mathbb{R} that is of order-type ω_1 in the inherited order. Suppose S were such a set. Observe that to the right of every element of S is an open interval disjoint from S . That is to

²I suspect that the the sequence $\langle f_\gamma(\zeta) : \gamma > \zeta \rangle$ of values given to ζ describe a nonmeasurable set. I have seen no proof of this, tho'. We needed AC to build it so it might well be nonmeasurable.

say \mathbb{R} is naturally partitioned into half-open intervals, and this partition is in 1-1 correspondence with S , each member of S being paired with the half-open interval of which it is the left endpoint. This partition can be injected into \mathbb{Q} by sending each piece to the first rational in it. So S was countable after all.

I have noticed that a surprising number of you use arguments involving countable choice.

One such argument says that, if there were a set X of reals of order-type ω_1 in the inherited order then each of the intersections $X \cap (n, n + 1]$ would be countable, meaning that X is a union of countably many countable sets and is therefore countable, contradicting the assumption that it is of length ω_1 and therefore of size \aleph_1 .

Using AC is bad practice even if AC is true. You don't want to use just any true fact that happens to be lying around: "God exists, so there is no order-preserving map from the second number class into the reals" doesn't quite cut it.

Some of you even managed to muck up the proof of two paragraphs above. OK, you send each countable ordinal to the open interval in \mathbb{R} as above. You then say: each interval contains a rational—which indeed it does—and then shut up shop and go home. That's not really good enough. The contradiction comes from having a function from a set of size \aleph_1 (the set of countable ordinals "the second number class") into a set of size \aleph_0 (the rationals). You can't stop until you have done it. You have to actually pick a rational from each of these intervals, so that you can send the countable ordinal in question to that rational. Which rational? With many of you it cost blood and threats of the rack to get you to say that the rationals have an ordering of length ω so you pick, from each interval, the first rational in that interval in the sense of that wellordering. Or you could look among the ordinals in that interval that have minimal denominator (there are only finitely many) and then plump for the smallest. Or the largest—who cares? Even after I had spelled this out, a lot of you clearly just thought I was barmy. Well, I'm not: what I was trying to get you to do was come up with a proof, not a nondeterministic add-warm-water-and-stir pseudoproof. That's Logic for you!

More temperately [calm down and breathe deeply, tf] what is going on here is that we want to prove that, were there *per impossibile* an object of the conjectured kind (to wit, an order-preserving injection from the set of countable ordinals into the reals) then there would be an object of a kind we know there cannot be, namely an injection of an uncountable set into a countable one. The proof must describe such a construction of an object of the second kind from an object of the first kind. One should never be *completely* satisfied with a nondeterministic construction if a deterministic construction is available.

If you want to think more about this have a look at www.dpmms.cam.ac.uk/~tf/fundamentalsequence.pdf

One of the things that this shows is that the quasiorder of linear order types (quasiordered by injective homomorphism) is not complete, or anything remotely like it: ω_1 and η (the order type of \mathbb{Q}) are distinct upper bounds for the second number class. ω_1 is a *minimal* upper bound but it is not the *minimum* upper bound, co's it ain't less than η . \mathfrak{c} (the order type of the reals) is an upper bound, but it is not a *minimal* upper bound; there is an infinite strictly descending sequence of upper bounds for the second number class all below \mathfrak{c} . (This is a theorem of Sierpinski, using a grubby diagonal argument powered by a wellordering of \mathbb{R} . I used to lecture it in my Part III lectures on WQO theory. It also shows its face in an Impossible Imre Question (question 14 sheet 2, 2015). Indeed this question remains undead, since it is Q14 on a sheet in at least one subsequent year. The IIQ is "Suppose $\langle X, \leq_X \rangle$ is a total order with no non-identity injective homomorphism Check the year

into itself. Must X be finite?”)

Actually it’s even worse than that: the quasiorder of linear order types isn’t even a poset, beco’s antisymmetry fails! (Consider $(0, 1)$ and $[0, 1]$.)

Apparently it’s an open question whether or not ω_1 and ω_1^* are the only minimal uncountable order types.

Question 10

Question 11

I have a student writing a Part III essay on this question. A lot there to think about. Very open-ended.

Question 12

Which of the following posets are complete?

(i) **The set of finite and cofinite subsets of \mathbb{N} , ordered by inclusion.**

It’s not a complete poset, since the set $\{\{1\}, \{1, 3\}, \{1, 3, 5\}, \dots\}$ does not have a supremum. That example also shows that it is not chain-complete.

(ii) **The set of independent subsets of a given vector space.**

The two elements $\{(1, 0), (0, 1)\}$ and $\{(1, 0), (1, 1)\}$ do not have a supremum, since any upper bound must include their union, and that is not linearly independent. However the collection of independent subsets of a vector space is of course *chain*-complete.

(iii) **The set of subspaces of a vector space, ordered by set-inclusion.**

This poset is complete. The supremum of any subset is the subspace spanned by the union of its elements.

(Observe that the *sup* and *inf* of this complete poset do not distribute. This is beco’s **inf** is “honest” [it’s just \cap] but **sup** is not: it’s sometimes bigger than \cup .)

Some of you, i notice, want to pick a basis for each subspace and then take the union of the bases. This is unnecessary, and indeed undesirable. The point is not that it uses the axiom of choice—tho’ it does—which is never a good idea if you can avoid it; the point is that it’s also a violation of the vector-space rule that you should always prefer basis-independent proofs wherever they are available.

Question 13

Some jottings, at this stage ...

If your finite boolean algebra \mathcal{B} wot you have in your hand is to be iso to a power set algebra, ask yourself: “power set of what?” The answer is of course the power set of its atoms. An atom is a minimal nonzero element. If \mathcal{B} is finite then every element of it has an atom \leq it. For suppose it didn’t: you could find another nonzero element strictly below it that wasn’t an atom, and you could build a finite descending chain as long as you please, in particular longer than the size of \mathcal{B} (which was finite, after all). (Notice that i have NOT used dependent choice!!)

Then you have to show that there is a 1-1 correspondence between elements of \mathcal{B} and the power set of the atoms. I.E., if $b \neq b'$ then there is an atom below one that is not below the other. If not, then $b' \text{ XOR } b$ has no atoms below it.

For the last part the answer is hidden in plain sight, as the first part of Q 12. It can't be a power set algebra because it's countably infinite, but it's definitely a boolean algebra. Another way in which infinite boolean algebras can differ from finite b.a.s is by not having atoms. A *regular open* set in a topological space is one that is equal to the interior of its closure (eg $(0, 1) \cup (1, 2)$ is open but not regular open). The regular open sets in a topological space always form a boolean algebra, and I think every boolean algebra can be shown to arise in this way (think of the order topology on the boolean algebra—I think!)

Question 14

These things are called *Lindenbaum Algebras*

Sheet 3

Question 1

Incorporating question 6

Use Zorn's Lemma to prove

- (i) that every partial ordering \leq_X on a set X can be extended to a total ordering of X ;
- (ii) that, for any two sets A and B , there exists either an injection $A \rightarrow B$ or an injection $B \rightarrow A$.

Discussion

(i) Give the set of partial orders on X the containment partial order as subsets of $X \times X$. The resulting partial order is chain-complete, since the union of a nested sequence of partial orders is still a partial order. (Incidentally do **not** make the mistake of thinking that your chains must be of the form $\{c_i : i \in \mathbb{N}\}$. It is true that hitherto every sequence you have ever seen has been indexed by the natural numbers—so that many of you think that ‘sequence’ just *means* ‘sequence indexed by \mathbb{N} ’—but in the big wide world outside Analysis there are chains too long to be indexed by \mathbb{N} and they might crop up here. The poset of countable ordinals is a rather in-your-face illustration! All its countable chains have suprema but its uncountable chains do not, and it has no maximal element. A poset all of whose countable chains have sups might not have a maximal element, because it might have *uncountable* chains.)

In fact the containment order is not only chain-complete but even *directed-complete*: every *directed* subset has a l.u.b. (A subset of a poset is *directed* if any two elements of it have a common upper bound in it.) By Zorn's lemma, it follows that, given any partial order \leq on X , there exist maximal partial orders extending it. Let \leq' be one such and let us establish that \leq' is a total order. Suppose it were not, and that we had $a, b \in X$ with $a \not\leq' b$ and $b \not\leq' a$. Then the relation \leq'' defined by $x \leq'' y$ iff $x \leq' y \vee (x \leq' a \wedge b \leq' y)$ is a partial order [easy to check that it is a partial order] properly extending \leq' —contradicting maximality of \leq' . ■

The poset of partial-orderings-under-inclusion is the obvious poset to use in (i), but one can also consider the set of total orders of subsets of X that are compatible with \leq_X —again, ordered by \subseteq . [You can guess what I mean by ‘compatible’ and it will do you no harm to write out a

formula that captures it.] There is no particular advantage to using this special partial order in this case, but a modification of it comes in handy in sheet 4.

Actually, thinking about the poset-of-partial-orderings... we have been thinking of the partial orderings as *sets of ordered pairs* ... in which case the order relation is plain old \subseteq , set inclusion. But it is probably worth making the point that you don't *have* to think of them as sets of ordered pairs. You can just think of them as relations (whatever *they* are!) and say that a partial order \leq is below (however you write that!) another partial order \leq' iff $(\forall xy)(x \leq y \rightarrow x \leq' y)$. You don't have to coerce everything in sight into being a set if you don't feel like it. Admittedly the coercion makes for a more uniform treatment but it doesn't actually shed any light on the proofs.

For (ii) you of course consider the chain-complete poset of partial bijections.

Question 2

Zorn's Lemma for countable posets.

You use the enumeration to ensure that the process of trying to reach a maximal element will succeed in finitely many steps.

Let $\langle X, \leq_X \rangle$ be a countable chain-complete poset. Enumerate X as $\langle x_i : i \in \mathbb{N} \rangle$. Build a \leq_X -chain the subscripts of whose elements form an $\leq_{\mathbb{N}}$ -increasing sequence. First one is x_0 , thereafter if the x -in-hand is maximal, then **HALT**; **else** plonk on the end that x which has $\leq_{\mathbb{N}}$ -minimal subscripts among those $x \geq_X$ the x -in-hand. If this doesn't **HALT** in finitely many steps the resulting chain has an upper bound and one obtains a contradiction by enquiring about the subscript on the upper bound.

I think the point of this question is to prepare you for a proof of ZL from AC. You want to show that a chain-complete poset $\langle X, \leq_X \rangle$ has a maximal element? Brutally wellorder X and use the technique of question 2.

Question 3

A general point worth making here. Suppose you are invited to use Zorn's lemma to prove the existence of a widget. To succeed you have to find a way of thinking of *widgets* as *dingbats that are maximal under \leq* . And the collection of dingbats, equipped with \leq (whatever that is!)—had better be a chain-complete poset. It's a classic exercise in pattern-matching.

Question 4

"And I never never use AC!"

"What, never?"

"No, never."

"What—*never*?"

"Well ... *hardly* ever."

W.S. Gilbert. H.M.S. Pinafore

Various places (probably—I didn't go to the lectures). One place it was certainly used is in proving that a sup of countably many countable ordinals is countable. Why does that need AC?

Another point (thank you, Dr Russell, for this piece of intelligence!) is that he used countable choice (in fact *dependent choice* aka DC) in showing that wellfoundedness of a relation R is equivalent to the nonexistence of infinite descending R -chains. Again, it would be good to explain to yourself why you need choice for that.

Question 5

I think this is bookwork. I am assuming that the theorem says that any inflationary function f from a chain-complete poset (we emphasise ‘set’) has a maximal element. You recursively define a function i from the ordinals into your set. $i(0)$ can be anything you like. Thereafter $i(\alpha + 1) =: f(i(\alpha))$ unless $i(\alpha)$ is fixed, and—for limit λ — $i(\lambda)$ is $\sup\{f(\alpha) : \alpha < \lambda\}$. You then appeal to Hartogs’.

I think that’s what Dr. Russell wants you to say, and it makes a lot of sense. It does sweep under the carpet some issues that the sharper-eyed among you might want to look at, such as “How do we know that defining functions by recursion on the ordinals in this way is OK?”. It is of course, but there is too much explanation required to fit it all within the confines of a Part II course.

And notice that the poset really has to be a set. If it is allowed to be a class then reflect that the power set function is an inflationary function on the chain-complete poset $\langle V, \subseteq \rangle$... but it has no fixed point beco’s of Cantor’s theorem.

Question 6

“Trichotomy of cardinals”. It implies AC. Easiest to infer AC in the form of the Wellordering Principle, the allegation that all sets can be wellordered. Suppose we wish to wellorder a set X . Hartogs tell us that there is a wellordered set too big to inject into X . But then, by trichotomy, X can be embedded into this set. But then of course we can copy back to X the wellordering of this set.

Some of you took the injunction to infer AC from trichotomy literally. I found myself reading things like “Let $\{X_i : i \in I\}$ be a family of pairwise disjoint nonempty sets. By assumption either $\{i\}$ injects into X_i or *vice versa* ...” The trouble with this method is that there may be lots of injections from $\{i\}$ into X_i and you need AC if you are to choose one, which begs the question³.

Question 7

I shouldn’t need to tell you this but ... don’t even *think* about using the axiom of choice. Do it, as they say, *synthetically*. Think of cardinals as equivalence classes of sets under equipollence.

Question 8

You all think you know that $|\mathbb{R}| = 2^{\aleph_0}$ and you’re right of course but finding a bijection between \mathbb{R} and $\mathcal{P}(\mathbb{N})$ is not an *absolute* doddle. It can do you no harm to track one down.

Jonathan Holmes has the cutest proof of this fact known to me (I have doctored this from his answer to an earlier sheet). Every real number has a unique representation as an ω -string of 0’s and 1’s *containing arbitrarily late 0s*. The italicised condition removes duplicate representations of dyadic rationals. Each such string corresponds to a set (of addresses where the string has 1s) whose

³Not often that you see that expression being used properly nowadays.

complement is infinite. How many subsets of \mathbb{N} are there whose complement in \mathbb{N} is infinite? Well, there are \aleph_0 subsets of \mathbb{N} that do not satisfy this condition, so we are looking at $\mathcal{P}(\mathbb{N})$ minus a countable set. You then use Bernstein's lemma to show that any such set has cardinality precisely 2^{\aleph_0} .

To return to Q8, when calculating that $2^{\aleph_0 \cdot 2^{\aleph_0}}$ you shouldn't say that $\aleph_0 \cdot 2^{\aleph_0} = 2^{\aleph_0}$ on the grounds that the product of two cardinals is the larger of the two; that uses AC, and you don't need it; instead you should say

$$2^{\aleph_0} \leq \aleph_0 \cdot 2^{\aleph_0} \leq 2^{\aleph_0} \cdot 2^{\aleph_0} = 2^{\aleph_0 + \aleph_0} = 2^{\aleph_0}$$

and appeal to Cantor-Bernstein.

The clever way to prove the second part is to observe that the function $F : (\mathbb{R} \rightarrow \mathbb{R}) \rightarrow (\mathbb{Q} \rightarrow \mathbb{R})$ that takes a continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$ and returns $f|_{\mathbb{Q}}$ (its restriction to \mathbb{Q}) is injective. Thus F injects the set of continuous functions $\mathbb{R} \rightarrow \mathbb{R}$ into the set of functions $\mathbb{Q} \rightarrow \mathbb{R}$. This latter set is of size $(2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0}$.

This illustrates a general phenomenon. If the set you are trying to compute the size of is naturally a subset of the n -times power set of \mathbb{N} or \mathbb{R} etc then its size will be an iterated exponent of \aleph_0 . However, sizes of quotients can be very complicated. How many wellorders are there of \mathbb{N} , *up to isomorphism*? The answer is \aleph_1 , which may or may not be equal to 2^{\aleph_0} . Contrast the two questions:

How many total orders of \mathbb{N} are there whose automorphism group is transitive on singletons? (A)

How many countable order types are there whose automorphism group is transitive on singletons? (B)

The answer to (A) is pretty obviously 2^{\aleph_0} . The set in (B) is a quotient of the set in (A) and calculating its size is a hard task.

Question 9

The first sentence contains an ellipsis. He doesn't really mean you to show that there is no surjection from \aleph_n onto \aleph_{n+1} beco's you don't know what sets these cardinals are. (In fact in both the von Neumann implementation and the Scott's trick implementation this allegation is, as it happens, true). What he means is that you should show that there is no surjection from **a set of size** \aleph_n onto **a set of size** \aleph_{n+1} . This is pretty easy, even if you aren't allowed choice. It follows from the simple observation that if there is a surjection $X \twoheadrightarrow Y$, and X can be wellordered, then there is an injection $Y \hookrightarrow X$.

The rest of the question is the **Jordan-König theorem**, which we will now prove.

THEOREM 1. *The Jordan-König theorem (AC):*

If $\langle A_i : i \in I \rangle$ and $\langle B_i : i \in I \rangle$ are families of sets such that $(\forall i \in I)(\text{there is no surjection } A_i \twoheadrightarrow B_i)$ then there is no surjection $\bigcup_{i \in I} A_i \twoheadrightarrow \prod_{i \in I} B_i$

Proof:

Consider any $f : \bigcup_{i \in I} A_i \rightarrow \prod_{i \in I} B_i$. We show that f is not onto. For each $i \in I$ let $f_i : A_i \rightarrow B_i$ be $\lambda x_{A_i}.(f(x))(i)$. f_i cannot be onto by hypothesis so (Remember we are using AC!) we pick n_i

to be a member of $B_i \setminus f_i "A_i$. Now we find that the function $\lambda i.n_i$ is not in the range of f , for o/w if $f(a) = \lambda i.n_i$ where $a \in A_i$ say, then $f_i(a) = (\lambda x.(f(x))(i))(a) = (f(a))(i) = \lambda i.n_i(i) = n_i$ contradicting choice of n_i . ■

This—full general—version needs choice⁴ but if you merely want to show that $2^{\aleph_0} \neq \aleph_\omega$ then you don't. (It will do you no harm to think about why not)

I think what Dr. Russell wants you to do is use that line of thinking to show that there is no surjection of $A \rightarrow (\mathbb{N} \rightarrow A)$ if $|A| = \aleph_\omega$... using the fact that \aleph_ω is the sup of the \aleph_n , with $n \in \mathbb{N}$.

Finally you might like to check your understanding of this situation by proving that 2^{\aleph_0} cannot be equal to \aleph_α if the cofinality of α is ω .

Question 10

(i): Fields of Characteristic 2

The language has $\Omega = \{+, \times, 0, 1\}$ with arities 2, 2, 0, 0 and $\Pi = \emptyset$. The theory can be described by the following axioms:

$$\begin{aligned} &(\forall x)(\forall y)(x + y = y + x) \\ &(\forall x)(\forall y)(\forall z)((x + y) + z = x + (y + z)) \\ &(\forall x)(x + 0 = 0) \\ &(\forall x)(\forall y)(x \times y = y \times x) \\ &(\forall x)(\forall y)(\forall z)((x \times y) \times z = x \times (y \times z)) \\ &(\forall x)(x \times 1 = x) \\ &(\forall x)(x \neq 0 \rightarrow (\exists y)(x \times y = 1)) \\ &(\forall x)(\forall y)(\forall z)(x \times (y + z) = x \times y + x \times z) \quad 1 + 1 = 0. \end{aligned}$$

(ii): Posets with no maximal element

The language has $\Omega = \emptyset$ and $\Pi = \{\leq\}$ with arity 2. The theory has the following axioms:

$$\begin{aligned} &(\forall x)(x \leq x) \\ &(\forall x)(\forall y)((x \leq y \wedge y \leq x) \rightarrow x = y) \\ &(\forall x)(\forall y)(\forall z)((x \leq y \wedge y \leq z) \rightarrow x \leq z) \\ &(\forall x)(\exists y)(x \leq y \wedge x \neq y) \end{aligned}$$

Be alert to the difference between **maximal** elements and **maximum** elements.

(iii): Bipartite graphs

This looks like an innocent question, but it has huge logical ramifications. Is a bipartite graph a graph that can be decorated with a two-colouring? Or is it a graph equipped with a two-colouring? these are two different kinds of things—different *signatures*

There are two correct answers, depending on which you mean.

(i) With a colour predicate:

The language has $\Omega = \emptyset$ and $\Pi = \{\sim, B\}$ of arities 2, 1. The theory has the following axioms:

⁴Is *equivalent* to choice indeed: think ... what happens if the product of nonempty sets is not reliably nonempty?

$$\begin{aligned}
& (\forall x)(\forall y)(x \sim y \longleftrightarrow y \sim x) \\
& (\forall x)(\forall y)(x \sim y \rightarrow (B(x) \wedge \neg B(y)) \vee (B(y) \wedge \neg B(x)))
\end{aligned}$$

(ii) But you can also do it without the colour predicate, by asserting that there are no cycles of odd length. This needs infinitely many axioms. You might like to prove that bipartite graphs cannot be finitely axiomatised in the language of graph theory: it's a useful compactness exercise of the kind that you might meet in an exam.

(iii) Actually there is a third correct answer which I hadn't considered, but which one of my students came up with. You could have a two-sorted language rather in the way that we might naturally have a two-sorted language for vector spaces. You have one set of variables for ranging over vertices, and another style of variable that ranges over colours. This is a much richer language and you can easily describe much more than just bipartite graphs. If you want a bipartite graph you have an axiom that says there are precisely two colours...

This method is of course extravagant, but the comparison between it and the method with a single colour predicate comes in useful later, with real vector spaces (part vii of this question). In part vii the analogue of method (iii) doesn't work: you have to do it by method one. But that's for later.

Finally, this question illustrates a lovely theorem of Kleene's: if T is a recursively axiomatised theory (has a recursive—or even merely a recursively enumerable—set of axioms) then you can add predicate symbols or function symbols to the language and you then get a finitely axiomatisable theory with the same theorems in the old language.

(iv): Algebraically Closed Fields

The language has $\Omega = \{+, \cdot, -, 0, 1\}$ with arities 2, 2, 1, 0, 0 and $\Pi = \emptyset$. The theory has the following axioms:

$$\begin{aligned}
& (\forall x)(\forall y)(x + y = y + x) \\
& (\forall x)(\forall y)(\forall z)((x + y) + z = x + (y + z)) \\
& (\forall x)(x + 0 = x) \\
& (\forall x)(x + (-x) = 0) \\
& (\forall x)(\forall y)(x \cdot y = y \cdot x) \\
& (\forall x)(\forall y)(\forall z)((x \cdot y) \cdot z = x \cdot (y \cdot z)) \\
& (\forall x)(x \cdot 1 = x) \\
& (\forall x)(x \neq 0 \rightarrow (\exists y)(x \cdot y = 1)) \\
& (\forall x)(\forall y)(\forall z)(x \cdot (y + z) = x \cdot y + x \cdot z) \\
& (\forall a_0) \dots (\forall a_n)(\exists x)(a_{n+1} \cdot x^{n+1} + a_n \cdot x^n + \dots + a_0 = 0).
\end{aligned}$$

where the last axiom is understood as an axiom scheme ranging over all positive integers n .

The point of this question is that there is a trap for the unwary: you can't reduce the scheme 'every polynomial of degree n has a root' to a single axiom because the degrees are not members of the field.

(v): Groups of Order 60

The language has $\Omega = \{\cdot, ^{-1}, 1, g_1, g_2, \dots, g_{60}\}$ with arities 2, 1, 0, 0, \dots 0 and $\Pi = \emptyset$. The theory can be axiomatised as follows:

$$\begin{aligned}
&(\forall x)(\forall y)(\forall z)(x \cdot (y \cdot z) = (x \cdot y) \cdot z) \\
&(\forall x)(x \cdot 1 = 1 \cdot x = x) \\
&(\forall x)(x \cdot x^{-1} = x^{-1} \cdot x = 1) \\
&(\forall x)(x = g_1 \vee x = g_2 \vee \dots \vee x = g_{60}) \\
&g_i \neq g_j \text{ for all } i \neq j \text{ (a scheme)}
\end{aligned}$$

That was QY's answer, but you can do it without the constants, which would be my preference.

(vi): Simple Groups of Order 60

You might think that the way in is to axiomatise the theory of simple groups and then add axioms (as above) to say there are precisely 60 elements. The trouble with that is that the theory of simple groups is not first order. (Exercise; *hint*: “abelian”). In general it's hard to tell whether-or-not a *prima facie* second-order theory is actually first-order after all. Consider the theory of rings with a unique maximal ideal, or the theory of posets wherein each element belongs to a unique maximal antichain.

So (back to (vi)) you have to exploit somehow the fact that it's only simple groups of order 60 that you are interested in.

You might think you can use group presentations to axiomatise the theory of simple groups of order 60, but it's less than completely straightforward. It's true that writing

$$\langle a^2 = b^3 = (ab)^5 = \mathbb{I} \rangle$$

in some sense captures A_5 , but writing

$$(\exists abc)(a \neq b \neq c \neq a \wedge a^2 = b^3 = (ab)^5 = \mathbb{I})$$

isn't enough by itself, since it appeals to the implicit convention that (i) no other equations hold; and (ii) everything is in the group generated by a and b . Neither of these assertions are first-order. However you might be able to get it to work if you additionally assert that there are precisely 60 elements. Dunno! Similarly It may be that it's enuff to say there are precisely 60 elements, and every element is of order 2, 3 or 5 and there are elements of all those orders; I don't know enough group theory to be sure.

The obvious way in, if you don't know that the only such group is A_5 (or don't intend to use the fact) is to say, of each n dividing 60, that whenever you pick up precisely n things they do not form a normal subgroup. In fact (one of my supervisees spotted this) it is sufficient to say ‘whenever i pick up 60 elements they do not form a normal subgroup’... the clever bit being that you don't say that the 60 things you pick up are all distinct!

However something has emerged recently which is that, in every finite simple group, every element is a commutator. My guess is that the converse is true too, namely that every group where every element is a commutator is simple. (The commutator subgroup is normal, so it had better be either trivial or the whole group). If that's true then you add to the axioms of Group Theory something to say that there are exactly 60 elements and

$$(\forall x)(\exists yz)(x = yzy^{-1}z^{-1})$$

Anyway the moral is that when you are trying to find a first-order axiomatisation of something that is obviously second-order you can—sometimes—cheat.

(vii): Real vector spaces

One's first thought is to axiomatise these vector spaces using a two-sorted first-order theory. One suite of variables (typically lower-case Greek letters) for ranging over scalars and another suite (typically lower case Roman letters) to range over vectors. One then adds the obvious axioms, starting with the axioms for a field of characteristic 0 and then adding stuff such as

- (1) $(\forall x)(\forall y)(\forall z)(x + (y + z) = (x + y) + z)$
- (2) $(\forall x)(\forall y)(x + y = y + x)$
- (3) $(\forall x)(x + 0 = x \wedge 0 + x = x)$
- (4) $(\forall x)(x + (-x) = 0 \wedge (-x) + x = 0)$
- (5) $(\forall xy)(\forall \lambda)(\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y)$
- (6) $(\forall x)(\forall \lambda, \lambda')((\lambda + \lambda') \cdot x = \lambda \cdot x + \lambda' \cdot x)$
- (7) $(\forall x)(\forall \lambda, \lambda')((\lambda \cdot \lambda') \cdot x = \lambda \cdot (\lambda' \cdot x))$
- (8) $(\forall)(0 \cdot x = 0)$
- (9) $(\forall)(1 \cdot x = x)$

where we “overload” ‘+’ to mean both the vector and scalar operations. Similarly ‘·’ covers both multiplication of scalars by scalars and multiplication of vectors by scalars.

If the field of scalars is finite (and therefore categorical—its theory has only one model) this works fine. However in our case the field is not categorical. A banal observation: if we try to set up a theory in a countable language then Skolem-Löwenheim will tell us that the theory has a countable model. In such a model there are only countably many scalars, so the model will be a vector space all right, but it won't be a vector space over \mathbb{R} ! OK, so you can blow it up by adding lots of constant symbols to ensure that the field of scalars is uncountable. But ensuring that the field is at least the size of the continuum isn't enough to ensure that it *is* the continuum. It might be something that looks like the reals but has infinitesimals. That is *fun* but it's not what we want and we have no space to discuss it here.

However the above theory is at least a point of departure. We modify it as follows. First we invent a constant symbol for every real. Next we write out lots of equations embodying the multiplication and addition tables for the scalars. Thus one adopts as axioms all true formulæ of the kind $a + b = c$, $a \cdot b = c$ for reals $a, b, c \dots$ (where the two operations are the scalar operations of course). Finally we replace any axiom involving scalar variables with all its (uncountably many!) instances—the result of replacing the scalar variables within it with constants in all possible ways. Notice that you now have no variables ranging over scalars.

Question 11

I'm assuming that the reader has discovered the back-and-forth construction. I'm not going to explain it here, coz it's best done interactively in real time. It's a slightly sexed-up (“symmetrical”) version of the construction that embeds every countable total order into \mathbb{Q} from the previous sheet wot i showed you in supervision: think of your countable total order as the members of \mathbb{N} written in a funny order, and then find homes for the natural numbers one by one. That's OK but sadly it isn't quite enough, coz it goes only one way. You might next think “Suppose I have two countable dense linear orders ... I can embed each in the other—so I can then use Cantor-Bernstein!” That doesn't work, beco's Cantor-Bernstein works for *cardinals* not for linear order types—they're far

too delicate. (After all, each of the two half-open intervals $(0, 1]$ and $[0, 1)$ embeds in the other but the two are not isomorphic.) So rather than build two embeddings separately, you *interleave* the two constructions in such a way that you construct a single isomorphism—a bijection.

Mind you, there actually *is* a version of Cantor-Bernstein for total orders, even tho' it is no use to us here. If A is iso to a terminal segment of B and B is iso to an initial segment of A then A and B are iso... Actually this is really a theorem about circular orders.

A follow-up thought...

Look at this once you've done sheet 4. Now that you have done ordinals and know what \aleph_1 is—the size of the set of countable ordinals—you might like to think about a generalisation of the fact that by a back-and-forth argument you can show that any two countable dense linear orders without endpoints are isomorphic. There is a theorem that says that any two dense linear orders of size \aleph_1 without endpoints are isomorphic (by a back-and-forth argument) as long as as they both satisfy a special extra condition.

What is that extra condition?

Question 12

Easy to show that the theory of fields of characteristic 0 is axiomatisable. Merely add the scheme

$$\underbrace{1 + 1 + \dots + 1}_{p \text{ times}} \neq 0 \quad \text{for each } p$$

to the field axioms.

Slightly harder to show that it is not *finitely* axiomatisable. We exploit the following trivial fact:⁵ Suppose T is a theory with an infinite axiomatisation A such that no finite subset of A axiomatises T . Then T has no finite axiomatisation. For suppose it did. Let ϕ be the conjunction of the finite set of axioms. We have $A \vdash \phi$. Then, by compactness, we have $A' \vdash \phi$ for some finite $A' \subseteq A$. But this, by hypothesis, we do not have. Observe that the above axiomatisation of the theory of fields of characteristic 0 is an infinite axiomatisation no finite subset of which suffices so we can exploit the trivial fact.

There is a temptation to think that if the theory of fields of characteristic 0 has a finite axiomatisation then it has one in which the field axioms are separately itemised, so that the remaining axioms can be conjoined into a single axiom which in effect says “the field is of characteristic 0”. Then you replace this axiom by its negation to obtain an axiomatisation of the theory of fields of positive characteristic, which of course is impossible. This can in fact be made to work, but it is not as straightforward as the proof i have just given. How can we be sure we can corral off the field axioms in this way? There is some work to do. Let our finite axiomatisation be the single formula ϕ . ϕ certainly implies the conjunction— F , say—of the field axioms. Now replace the single axiom ϕ with the two axioms $F \rightarrow \phi$ and F .

Are we now home and hosed? The candidate theory of fields of positive characteristic we obtain will be the field axioms F plus the negation of the remaining axiom $F \rightarrow \phi$. This negation is $F \wedge \neg\phi$, so this amounts to adding $\neg\phi$ as an axiom. Clearly no model \mathfrak{M} of $F \wedge \neg\phi$ can be a model of ϕ so \mathfrak{M} must be a field [beco's $\mathfrak{M} \models F$] and a field of positive characteristic. Converse? Let \mathfrak{M}

⁵I *know* it is trivial beco's i worked this out for myself when i was a mere philosophy student... a much lower lifeform than you, Dear Reader!

be a field of positive characteristic. It's a model of F , because it's a field, but it can't be a model of ϕ because it isn't of characteristic 0. So $\{F, \neg\phi\}$ would be an axiomatisation of the theory of fields of positive characteristic [which we know to be impossible] so there really is no such ϕ .

Question 13

A group with an element of infinite order.

The language has $\Omega = (\cdot, ^{-1}, 1, g)$ with arities 2, 1, 0, 0 and $\Pi = \emptyset$. The theory can be described by the following axioms:

$$\begin{aligned} &(\forall x)(\forall y)(\forall z)(x \cdot (y \cdot z) = (x \cdot y) \cdot z) \\ &(\forall x)(x \cdot 1 = 1 \cdot x = x) \\ &(\forall x)(x \cdot x^{-1} = x^{-1} \cdot x = 1) \\ &\underbrace{g \cdot g \cdot \dots \cdot g}_{n \text{ times}} \neq 1 \text{ (for each } n \in \mathbb{N}) \end{aligned}$$

Can this be done purely in the language of groups? The answer the question-setter wants is 'no' and he is obviously correct, as we will see. However, Prof. Leader (for it was he) is a mere mortal (tho' he might not appear to be, on cursory inspection) and the question contains a mistake. Since it is possible to axiomatise group theory just in the language with a single binary function symbol, you can go ahead and do it that way and—since you now no longer need the symbol ' e ' to denote the unit—you can recycle that symbol to denote the element of infinite order! But that's cheating; clearly the student who did this is destined for a life of crime.

There now follows a proof of the impossibility of doing this in the language of groups, reconstructed from a conversation I had with Prof. Leader.

The key is to find two groups one of which has an element of infinite order and the other does not, and yet the two groups are elementarily equivalent (indistinguishable by first-order expressions). To this end consider a group with elements of arbitrarily large finite order but no elements of infinite order. The group $\text{FSym}(\mathbb{N})$ of permutations of \mathbb{N} that move only finitely many things will do nicely. Now consider the theory $T = \text{Th}(\text{FSym}(\mathbb{N}))$ consisting of all the expressions in the language of group theory that hold in this group. This theory might not have a decidable set of axioms, but it doesn't matter. What *does* matter—indeed is absolutely crucial—is that it is a **complete** theory. We now add a constant g to the language, and the obvious axioms $g^n \neq e$, for all $n \in \mathbb{N}$. Call the resulting theory T' . T' is clearly consistent by compactness and must have a model, which will be a group, call it G . G is a model of the complete theory $\text{Th}(\text{FSym}(\mathbb{N}))$ and is therefore elementarily equivalent to $\text{FSym}(\mathbb{N})$. But G has an element of infinite order and $\text{FSym}(\mathbb{N})$ does not.

It doesn't much matter that we took our group to be $\text{FSym}(\mathbb{N})$. Any group with elements of arbitrarily large finite order but none of infinite order will do.

This works, and it's very pretty, but it's a bit *ad hoc*, and it's certainly not the kind of example that would naturally occur to people with your conditioning history. Nathan Bowler points out to me that the additive group of the rationals mod 1 ("the rational circle") has no element of infinite order (p/q is of order q) but the reals mod 1 ("the real circle") has elements of infinite order. My guess is that these two groups are elementarily equivalent, and indeed that the inclusion embedding is elementary. By this we mean that, for any expression $\phi(\vec{x})$ in the language of groups, if $\phi(\vec{p})$

holds of some tuple \vec{p} in the additive group of the rationals mod 1, then it holds of the same tuple of rationals in the bigger group of reals mod 1. I might get round to writing out a proof. If it works (and i'm not making any promises) it would be a nicer proof than Prof Leader's (although it's much more involved) beco's it is an introduction to a new technique.

Question 14

The theory T has the following axioms:

$$\begin{aligned} &(\forall x)(\forall y)(f(x) = f(y) \rightarrow x = y) \\ &(\forall y)(\exists x)(f(x) = y) \\ &(\forall x)(\underbrace{f(f(f \cdots (x) \cdots))}_{n \text{ times}}) \neq x \text{ (for each } n \in \mathbb{N}) \end{aligned}$$

Any countable model \mathfrak{M} of T is a disjoint union of at most countably many f -cycles, all of which are of the form $\{\dots f^{-2}(x), f^{-1}(x), x, f(x), f^2(x), \dots\}$ for some x .

Imagine you are living in a world where there is nothing going on other than lots of points joined together by f edges, and all you can ever do is move along f edges (in either direction) from one point to another. What do you discover? By the end of time you have discovered that you are living on a copy⁶ of \mathbb{Z} . And that's *all* you have discovered: if the model contains another copy of the \mathbb{Z} -gon that you could have been on you never learn this fact. There is no way, in the given language, of saying that two vertices lie on distinct \mathbb{Z} -gons.

This is an informal picture and is definitely not a proof, but it might lead us to one.

I *think* that the model consisting of a single copy of \mathbb{Z} is what they call a **prime model**: it injects elementarily into all models of T . Presumably we use quantifier-elimination.

This could serve as an introduction to *Ehrenfeucht Games* but i can't go into that sort of detail here.

But there is a proof using only techniques available to you. (There must be, since this question isn't starred.) You observe that, altho' T can have nonisomorphic *countable* models (one, two or many copies of \mathbb{Z}), all its models of size 2^{\aleph_0} are isomorphic. This may not be immediately obvious. If \mathfrak{M}_1 and \mathfrak{M}_2 are two models both of size 2^{\aleph_0} then they both consist of 2^{\aleph_0} \mathbb{Z} -gons. (A detailed proof of this fact needs a little bit of AC but i'll spare you the details). So there is a bijection between the (set of) \mathbb{Z} -gons-in- \mathfrak{M}_1 and the (set-of) \mathbb{Z} -gons-in- \mathfrak{M}_2 . This isn't *quite* a bijection between \mathfrak{M}_1 and \mathfrak{M}_2 , but we are nearly there. All we have to do is pick, for each pair of a- \mathbb{Z} -gon-in- \mathfrak{M}_1 -with-a- \mathbb{Z} -gon-in- \mathfrak{M}_2 , a digraph isomorphism between the two \mathbb{Z} -gons, and take the union of all those isomorphisms. This union will be an isomorphism between \mathfrak{M}_1 and \mathfrak{M}_2 . If T were not complete we would be able to find ϕ such that $T \cup \{\phi\}$ and $T \cup \{\neg\phi\}$ were both consistent. Add 2^{\aleph_0} constants and deduce (by compactness) that $T \cup \{\phi\}$ and $T \cup \{\neg\phi\}$ both have models of size at least 2^{\aleph_0} . Indeed (by downward Skolem-Löwenheim) they must both have models of size *precisely* 2^{\aleph_0} . These models would have to be nonisomorphic beco's one of them believes ϕ and the other believes $\neg\phi$. But they are both models of T so they are isomorphic.

Instead of 2^{\aleph_0} one can use \aleph_1 . Students would be unlikely to try doing it that way beco's \aleph_1 is a mysterious phobic object for them. But it works better. In particular one does not need AC,

⁶Actually it's not really \mathbb{Z} beco's \mathbb{Z} has additive and multiplicative structure, which this thing hasn't. It's really just a digraph. One might call it the **\mathbb{Z} -gon**.

at least not after the use of AC to prove upward Skolemheim to show that there is a model of size \aleph_1 . What does a model of T of size \aleph_1 look like? Lots of copies of \mathbb{Z} of course, but precisely how many? The set of copies of \mathbb{Z} is a surjective image of a set of size \aleph_1 and so is of cardinality $\leq \aleph_1$. The copies of \mathbb{Z} have a global wellordering, so the size of their union (which is \aleph_1) is \aleph_0 times something; that something can only be \aleph_1 .

Sometimes students can be soooo annoying. The point of this question (as you have probably guessed by now) is to direct your attention to theories that are categorical in some *uncountable* cardinal. However there is a way of answering this question that doesn't exploit this possibility, and some of you found it. That was not in the script at all. Grrr! Suppose $T \not\models \phi$ and $T \not\models \neg\phi$. Add countably many constants to the language of T , and add axioms to $T \cup \{\phi\}$ and to $T \cup \{\neg\phi\}$ to say that the denotations of these constants all belong to different \mathbb{Z} -gons. These two theories (call them T_1 and T_2) both have countable models by downward Skolemheim. What can countable models of these theories look like? They must consist of precisely \aleph_0 \mathbb{Z} -gons infinitely many of which have a distinguished element in each \mathbb{Z} -gon. There's no way of compelling every \mathbb{Z} -gon to have a distinguished element, so this doesn't completely wrap things up for us. However, you weren't supposed to do it that way anyway!

Question 15⁺

This is an interesting and important fact: logic has a lot to say about field theory. However i can't see any way of answering this question with the tools you have been given so far.

1.1 Sheet 4

Question 1

Challenge: deduce the axiom of empty set from the axiom of infinity and the axiom scheme of separation. Does Infinity imply empty set by itself?

The morally correct way to do this is to observe that the axiom of infinity has the form “there is a set with a special property”. If there is even one set, then—as long as we have separation—there will be an empty set, since the subsets consisting of all those elements of the set that are not equal to themselves will be a set by separation.

Now the axiom of infinity can also come in the form “There is a successor set”, or

$$(\exists x)(\emptyset \in x \wedge (\forall y)(y \in x \rightarrow y \cup \{y\} \in x))$$

In the presence of the axiom scheme of replacement this can be deduced from the bare assertion there is an infinite set (a set not the same size as any proper subset of itself) but the axiom is often taken in this more specialised form because it makes it easy to give immediately an implementation of arithmetic. Fair enough. However, this muddies the waters slightly, in that it enables us to give a different proof of the existence of the empty set. People sometimes say that the axiom of infinity *presupposes* the existence of the empty set, but that's not quite right. Let's get this 100% straight. The axiom in the form “there is a successor set” says

$$(\exists x)((\exists e \in x)(\forall w)(w \notin e) \wedge (\forall y)(y \in x \rightarrow y \cup \{y\} \in x)) \quad (1.2)$$

I have written out the ‘ $\emptyset \in x$ ’ bit in primitive notation so we can be sure that there are no tricks being played.

The expression (2.2) is of the form

$$(\exists x)(p \wedge F(x))$$

where p is $(\exists e \in x)(\forall w)(w \notin e)$ and $F(x)$ is $(\forall y)(y \in x \rightarrow y \cup \{y\} \in x)$. Anything of the form $(\exists x)(p \wedge F(x))$ is going to imply $(\exists x)p$, namely

$$(\exists x)(\exists e \in x)(\forall w)(w \notin e)$$

whence

$$(\exists e)(\forall w)(w \notin e)$$

which says that there is an empty set, which is what we wanted.

Does separation imply empty set? No, for the annoying (but important) reason that the empty model trivially satisfies separation but does not contain the empty set. This is a level nought independence proof: exhibit a countermodel!

Actually some of you have made the connection between this question and the point (made in lectures) that we do not admit the empty model. This question of whether or not we admit the empty model is a vexed one that i have striven not to think about for most of my working life—and with moderate success. Aristotle did not accept empty models, since he thought that “All A are B ” implied “Some A are B ”. Recently (i.e., within my lifetime) people (Prof Hyland among them) have started saying that perhaps it would be a good idea to admit empty models. If you want to pursue this, google ‘Free Logic’.

My current thinking is that since we evidently cannot deduce the empty set axiom from the separation scheme, and since the only countermodel (= a model that shows that you cannot perform that deduction) is the empty model, then if we want to preserve the completeness theorem, well, we’d better accept empty models. However, as i say, this is something i have tried not to think about!

Question 2

Question 2:1

... deducing the axiom (scheme) of separation from the axiom (scheme) of replacement.

If replacement allows you to use *partial* functions it’s easy. If you are only allowed *total* functions then you want Phil Connell’s trick (tidied up by me to make it constructive):

Define $f(x)$ to be $\{y : y = x \wedge \phi(y)\}$. This has the effect that f sends to their singletons the things you want to keep, and sends everything else to the empty set. Then $\bigcup f “W$ is $\{x \in W : \phi(x)\}$. Observe that this is constructive.

The other way (preferable in certain circumstances) is to say: *either* there is nothing in W which has ϕ (in which case the set we want is the empty set, and we have an axiom for that) *or* there is an $x \in W$ s.t. $\phi(x)$. For any such x we can define a function f which sends $y \in W$ to y as long as $\phi(y)$, and sends y to x o/w.

What is there to choose between these two proofs? Phil Connell's proof uses the axiom of sumset, but the second method uses excluded middle. (It uses it *twice*; once in the case split, and again in the second of the two cases, testing whether or not $\phi(y)$).

Question 2:2

Two applications of power set to \emptyset gives you $\{\{\emptyset\}, \emptyset\}$ which we then whack with the function class

$$(u = \{\emptyset\} \wedge v = x) \vee (u = \emptyset \wedge v = y)$$

which will give us the pair $\{x, y\}$.

Question 3

“Write down sentences in the language of set theory to express the assertions that, for any two sets x and y , the product $x \times y$ and the set y^x of all functions from x to y exist.”

The phrase “You may assume that your pairs are Wiener-Kuratowski” has been omitted from this year's version of this question. For the moment i am going to pretend that it wasn't omitted!

If you use Wiener-Kuratowski pairs then $\langle x, y \rangle = \{\{x\}, \{x, y\}\}$ and is a subset of $\mathcal{P}^2(\{x, y\})$. Similarly $x \times y$ is a subset of the power set a couple of times of $x \cup y = \bigcup \{x, y\}$. Clearly the set of functions from x to y can be obtained in the same way. What if you want to establish that

these things are sets without knowing what your pairing function is? (Which is evidently what Dr. Russell is expecting of you) Imagine the following situation: i want $X \times Y$ and i know that there is a set-theoretic construct $\langle x, y \rangle$, tho' i don't know what it is and i'm not allowed to assume anything other than that it is there and is available. We do the following: fix $y \in Y$ and consider the function class that sends x to $\langle x, y \rangle$. The image of X in this function exists by replacement and it is of course $X \times \{y\}$. So $X \times \{y\}$ exists for all y . Now consider the function class that sends y to $X \times \{y\}$. The image of Y in this function exists by replacement and is $X \times Y$.

So: if we have replacement we can prove that $X \times Y$ exists *whatever implementation of pairing-with-unpairing we use*. You might like to prove the converse: if $X \times Y$ always exists for all implementations of pairing-with-unpairing then replacement follows.

Question 4

I think you know by now that a Von Neumann ordinal is a transitive set wellordered by \in . “ x is transitive” is $\bigcup x \subseteq x$ and “ x is wellordered by \in ” can be captured in various ways. If we exploit the axiom of foundation (which is standard practice) then it is enough to say that

$$(\forall y, z \in x)(y \in z \vee y = z \vee z \in y)$$

because that will force the restriction $\in|_x$ of \in to x to be transitive. (It may be worth spelling this out... suppose $y, z, u \in x$, and $y \in z \in u$. By our assumption we must have $y \in u \vee y = u \vee u \in y$. But the second and third disjuncts contradict foundation.)

$$\left(\bigcup x \subseteq x\right) \wedge (\forall y, z \in x)(y \in z \vee y = z \vee z \in y) \tag{***}$$

I'm not quite sure what he wants as an answer to “What should the Von Neumann ordinal ω^2 be?”

It's a Von Neumann ordinal (and we know how to say that something is a Von Neumann ordinal). What else can we say, that might distinguish it? Well, consider those “limit” elements of x that are not successors ($y \cup \{y\}$) of any other member of x . This property is captured by $\text{limit}(y)$ iff $(\forall z \in x)(y \neq z \cup \{z\})$. Observe that every $y \in x$ such that $\text{limit}(y)$ has a maximal $\text{limit}(z)$ below it (if it has any at all). We can capture that ... and i am going to leave the details to you.

Naturally this can be written down in the language of set theory. It may be worth thinking about quite which Von Neumann ordinals can be uniquely characterised by formulæ of the language of set theory in this way. Clearly only countably many. There is some interesting mathematics here which i should probably think about, tho' i appreciate that you have other fish to fry. However it may be worth making the point that whether or not a formula in the language of set theory uniquely identifies a Von Neumann ordinal depends not just on the formula but also on the axioms of that theory. More axioms, more ordinals get identified. For example, the formula (***) provided above for “ x is a von Neumann ordinal” works only if you have the axiom of foundation—which mostly one does, of course. The point is that if you add further axioms then other formulæ which *didn't* denote ordinals suddenly do: (***) is a case in point; without foundation it doesn't characterise von Neumann ordinals.

Question 5

Neither direction works.

$\in\{\{\emptyset\}\}$ is the empty relation, and therefore transitive, but $\{\{\emptyset\}\}$ is not a transitive set.

For the other direction, consider

$\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$ is a transitive set, but $\in\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$ is not a transitive relation.

I think the danger in this second part is to reason “Suppose $z \in y \in x$; then, since $\in x$ is transitive, we infer $z \in x$ ”. This doesn't work. It assumes that $\in x \cup \{x\}$ is transitive which sounds the same but isn't.

Question 6

Read the chapter in Logic, Induction and sets.

Question 7

First part: 7 and 8. He's obviously taking natural numbers to be the corresponding Von Neumann ordinals and the rank of a von Neumann ordinal is itself⁷.

The second part of this question makes several points. One of them is the point that there are lots of ways of implementing \mathbb{Z} , \mathbb{Q} , \mathbb{R} etc as sets; another is that—mathematically at least—it doesn't much matter which one you use. The other is to get you to do some set-theoretic calculations—computing the ranks of particular sets.

One should start with a warning : the (set-theoretic) rank of a set equipped with an ordering cannot be computed from the order-type of the ordering: it's a property of the set, not of any

⁷Actually, here i should show my true self; the rank of a set is an ordinal, and ordinals are not sets, they are numbers. So what I should really say is the rank of a von Neumann ordinal is not itself, but the ordinal of which it is ... the von Neumann ordinal

ordering of it. And again, it's nothing to do with cardinality either, or very little. There are small sets (*singletons* indeed) of arbitrarily high rank.

So the rank of a mathematical object implemented as a set is not a mathematical invariant of that object. This comes as a surprise to many students, so it is worth making a big song-and-dance about it. Admittedly it is true that the *minimum possible* rank of an implementation of that object is a mathematical invariant of that object [I think Prof Leader sometimes calls it “essential rank”] but it's a curiously uninteresting one, being controlled entirely by cardinality. You cannot implement \mathbb{R} as an object of rank ω beco's there are too few things of lower rank for all the reals to be implemented by those things of lower rank. There are uncountably many reals but only countably many things of finite rank. This cardinality consideration is the only constraint.

Actually in parts of the set theoretic literature reals are taken to be functions from \mathbb{N} to \mathbb{N} , things sometimes called *set theorists' reals*. They have rank ω —which is best possible, for the reason given above. (They're something to do with continued fractions.)

So: pick an implementation, and compute the ranks of the sets you end up with. For bonus points, pick more than one implementation, and compute all of them! If you know what p -adic numbers are, compute their rank too. (The p -adics are the completion of \mathbb{Q} w.r.t the p -adic metric. How are you to think of the completion set-theoretically?)

Question 8

The key to this question is induction, both structural and wellfounded.

One direction is easy: you prove by induction on n that everything in V_n is in HF . For the other direction you have to use \in -induction to show $HF \subseteq V_\omega$. The property $\phi(x)$ you prove by \in -induction is “ $x \in HF \rightarrow x \in V_\omega$ ”.

If foundation fails then potentially a Quine atom⁸ is a counterexample to the inclusion.

This is quite a good question to look at. You have two inclusions to prove, one in each direction. To prove $V_\omega \subseteq HF$ you do an ordinary mathematical induction—on rank; to prove $HF \subseteq V_\omega$ you do an \in -induction. A nice idiomatic illustration. Beco's of the possibility of finding Quine atoms in HF (which violate foundation) you *have* to use foundation to prove that $HF \subseteq V_\omega$.

HF is more usually known as V_ω . It is a model of all the axioms of ZFC except infinity.

Codicil to questions 7–9

Take care when asking yourself whether or not an axiom is true in a structure. Yer typical set theoretic axiom states that the universe is closed under some operation (as it might be power set, or sumset). **Saying that a structure is a model for that axiom is not the same as saying that it's closed under the corresponding operation.**

Question 9

Which axioms of ZF hold in $V_{\omega+\omega}$?

Discussion

All of them except replacement. Consider the function class $n \mapsto V_{\omega+n}$. Replacement would make the image of \mathbb{N} in this function class—namely $\{V_{\omega+n} : n \in \mathbb{N}\}$ —into a set of the model, and it

⁸A Quine atom is a set $x = \{x\}$.

can't be, beco's it is of rank $\omega + \omega$.

Observe that the fact that you can find a subset of \mathbb{R} that can be wellordered to length $\omega + \omega$ has the following ramifications. If replacement held in $V_{\omega+\omega}$ then $V_{\omega+\omega}$ would contain, for every wellordering in $V_{\omega+\omega}$, the corresponding von Neumann ordinal. It is easy to check that the rank of the von Neumann ordinal α is α itself, which means that $V_{\omega+\omega}$ cannot contain any ordinal from $\omega + \omega$ onwards. So replacement fails. In general V_α will contain at least some wellorderings that are far too long for their von Neumann ordinals to be in V_α . It happens only rarely that α “catches up” with the ordinals in V_α .

In contrast, the collection H_κ of sets hereditarily of size less than κ is practically guaranteed to be a model of replacement, as follows. Suppose $X \in H_\kappa$, and $f : H_\kappa \rightarrow H_\kappa$. Then $f''X$ is a subset of H_κ . How big is it? It's a surjective image of thing of size $< \kappa$. We want it to be of size $< \kappa$ itself. So all we need is a surjective image of something of size less than κ is itself of size $< \kappa$. This is certainly true if κ is an aleph, and even in many cases when it isn't. So certainly if κ is an aleph then H_κ is a model of replacement.

Question 10

This is all about absolute properties *versus* non-absolute properties. Ha Thu Nguyen gives a very simple illustration ... If $\mathfrak{M} = \langle M, \in \rangle$ is a model of ZF then it doesn't think that its carrier set M is a set, but we can see from outside that it is. The thing that \mathfrak{M} believes to be the power set of \mathbb{N} is, indeed, countable seen from outside; however \mathfrak{M} does not know of any bijection between that set and the set it believes to be \mathbb{N} . Tim Talbot puts it very well: the countable model is a *Tardis*!

Question 11

For the first part: compactness! (This should be obvious; if it isn't then seek help). For the second part, the descending sequence is not a set of the model and is therefore not a counterexample to wellfoundedness of the order relation on ordinals.

On further reflection (Thank you Tim Cooper!) it occurs to me that it might be helpful to say a bit about quite *why* the descending sequence of ordinals is not a set of the model (tho' i think this may be going beyond what Dr Russell was expecting of you). One could be forgiven for thinking: “surely the sequence is a set ... we take the image of \mathbb{N} in the function $n \mapsto \alpha_n$!”. The point is that this function is not a function class (even of) the expanded language with the new constants. (The kind of nitpicking one has to do in order to see precisely what is going on marks out the logician from the ordinary mathmo.) We added countably many constants to the language of set theory. We, the user, can see the structure on the set of constants—we can see the numerical subscript in each constant. However the language itself is oblivious to such details. As far as it is concerned they are just constants *and have no internal structure*. The numerical subscripts are not visible, detachable, parts of the constants that can be replaced by variables.

It's a subtle annoying point but it matters.

Question 12

A union of countably many countable sets cannot have size \aleph_2 .

I think of this as a rather fun question, tho' discerning people like Prof Leader and Dr Loewe think it's hard. I think it's not so much hard as really quite scary, since \aleph_2 is a radically unfamiliar

object. I would bet good money that you have never seen an object of this size ever before. I would even bet good money that you cannot give me an example of a thing of this size.

Ndevertheless it's an idiomatic piece of bare-hands set theoretic manipulation, and you will find it satisfying.

It's easier if you attempt to prove the rather more general:

if every A_α is of cardinality \aleph_γ at most, **and** $\bigcup_{\alpha < \omega_\beta} A_\alpha$ can be wellordered
then $|\bigcup_{\alpha < \omega_\beta} A_\alpha| \leq \max\{\aleph_\beta, \aleph_\gamma\}.$

The point being that if $\bigcup_{\alpha < \omega_\beta} A_\alpha$ is wellordered then you can use the restrictions of that wellordering to the various A_i to exploit the fact that $\aleph_\gamma \cdot \aleph_\gamma = \aleph_\gamma$.

If you want a concrete actual set of size \aleph_2 you can think of the Von Neumann ordinal ω (or the set I_{ω_2} or ordinals below ω_2 which is actually the same thing). Then you consider what happens if you express this set as a union of countably many countable sets.

There are analogues of this that you can prove by the same method: a union of countably many finite sets cannot have size \aleph_1 ; a union of \aleph_1 countable sets cannot be of size \aleph_2 .

For all of these you need that $(\aleph_\alpha)^2 = \aleph_\alpha$. Some of you got your wires crossed on being told that AC is equivalent to $(\forall \alpha)(\alpha = \alpha^2)$ ("cardinals are idempotent") and came away with the impression that proving that $(\aleph_\alpha)^2 = \aleph_\alpha$ needs the axiom of choice. It doesn't. There are two things going on: (i) AC is equivalent to "every cardinal is an aleph"; (ii) "alephs are idempotent" you prove directly, and without AC. I think Dr R proved it in lectures. In case he didn't there is a proof in my lecture notes for this course from 2016, linked from my home page on https://www.dpmms.cam.ac.uk/~tf/cam_only/partiilectures2016.pdf.

Question 13

Is every countable model of first-order Peano arithmetic isomorphic to the set of natural numbers?

QY sez: "No: adjoin to the language a constant c and adjoin to the axioms of Peano arithmetic the sentences $0 < c$, $s(0) < c$, $s(s(0)) < c$, ... to obtain a new theory S . Each finite subset of S has a model, so by compactness S has a model, which is of course infinite. By downward Löwenheim-Skolem, it has a countable model \mathfrak{M} . In \mathfrak{M} there is an element c which is greater than 0, $S(0)$, $S(S(0))$... but there is no such element in the standard model \mathbb{N} , so \mathfrak{M} is a nonstandard countable model of Peano arithmetic."

Thanks for this QY, but classroom experience teaches me not leave it at that. Very well, so we have a model of arithmetic with an extra element. But it doesn't stop there. PA proves a whole lot of theorems saying that \mathbb{N} is closed under a lot of operations: $x \mapsto x^2$, $x \mapsto \lceil 22x/7 \rceil$, $x \mapsto \lceil \sqrt{x} \rceil$ and so on. It is probably quite helpful to think of our model as something containing 0 and c and *generated by them*. At its most basic it is a theorem of the arithmetic of \mathbb{N} , after all, that every number has a successor—and that every nonzero number has a predecessor—so we must have $c+1$ and $c-1$. This leads us to the conclusion that c belongs to a copy of \mathbb{Z} stuck on the end of \mathbb{N} . Only one copy...? What about $\lceil 22c/7 \rceil$, $\lceil 355c/133 \rceil$...? In fact we have a copy of \mathbb{Z} for every rational!

Question 11 on the preceding sheet tells us that there is only one countably infinite dense linear order without endpoints (up to isomorphism). In this context Q11 tells us that all countable nonstandard models of PA are isomorphic **as ordered sets**. (Not as structures for the language of arithmetic, of course. Check that you are happy with this). So every countable nonstandard model of PA has order type $\mathbb{N} + \mathbb{Q} \cdot \mathbb{Z}$. You might think that you get *more* than \mathbb{Q} copies of \mathbb{Z} beco's of $\lceil \sqrt{c} \rceil$ but—as noted above, \mathbb{Q} is a maximal countable linear order type so you don't get any further copies of \mathbb{Z} by considering $\lceil \sqrt{c} \rceil$. Of course they aren't all isomorphic *as structures for $+$ and \times* —beco's arithmetic is incomplete.

I have just learnt the curious fact that every countable nonstandard model of PA is isomorphic to a proper initial segment of itself!

One point one sometimes has to make in this connection is that these wild and woolly things—the nonstandard naturals—living in the desolate marches beyond the standard naturals are absolutely **not** the same wild and woolly things living in the desolate marches beyond ω , namely the countable ordinals. This mistaken identification is a common consequence of over-enthusiastic fault-tolerant pattern matching by beginners. (Nor are either of those tribes of wild and woolly things the same as the cardinals of infinite Dedekind-finite sets!!!)

Question 14

Yes (obviously) and no (equally obviously). This question is put in as a reality check.

Question 15*

“What is going on here?” one wonders; what is at stake? The point is that once one has gadgetry for **lists** one can give *direct* definitions for functions defined by recursion over \mathbb{N} . (Notice that recursive definitions are *prima facie* circular.) It's quite a lot of work and often not spelled out properly, so when i lectured Part III in 2016 i went through this in some detail. Have a look at section 3.1.2 of https://www.dpmms.cam.ac.uk/~tf/cam_only/partiiicomputability.pdf.

...but only if your curiosity is piqued by the thought that it really matters that one can obtain *noncircular* definitions from circular (recursive) definitions, and that it matters how one does it. Most of you will be quite happy merely to be reassured that it can be done and won't worry about the details.

Chapter 2

Professor Leader's Example Sheets for 2019

Sheet 1

Question 1

Duplicates Dr Russell's sheet 1 q 1

Question 2

[PTJ sez (inter alia) *The fact that $\{\neg p\} \vdash (p \rightarrow q)$ is needed in the proof of the Completeness Theorem.*]

QY supplies this proof.

By the deduction theorem, it suffices to show that $\perp \vdash q$. The following is a proof:

t_1	\perp	(in S)
t_2	$\perp \rightarrow ((q \rightarrow \perp) \rightarrow \perp)$	K
t_3	$(q \rightarrow \perp) \rightarrow \perp$	(modus ponens from t_1, t_2)
t_4	$((q \rightarrow \perp) \rightarrow \perp) \rightarrow q$	(axiom 3)
t_5	q	(modus ponens from t_3, t_4)

Then by the proof of the deduction theorem, the following is a proof that $\perp \rightarrow q$:

1	$\perp \rightarrow (\perp \rightarrow \perp)$	K
2	$\perp \rightarrow ((\perp \rightarrow \perp) \rightarrow \perp)$	K
3	$(\perp \rightarrow ((\perp \rightarrow \perp) \rightarrow \perp)) \rightarrow ((\perp \rightarrow (\perp \rightarrow \perp)) \rightarrow (\perp \rightarrow \perp))$	S
4	$(\perp \rightarrow (\perp \rightarrow \perp)) \rightarrow (\perp \rightarrow \perp)$	(modus ponens from 2, 3)
5	$\perp \rightarrow t_1$	(modus ponens from 1, 4)
6	t_2	K
7	$t_2 \rightarrow (\perp \rightarrow t_2)$	K
8	$\perp \rightarrow t_2$	(modus ponens from 6, 7)
9	$(\perp \rightarrow t_2) \rightarrow ((\perp \rightarrow t_1) \rightarrow (\perp \rightarrow t_3))$	S
10	$(\perp \rightarrow t_1) \rightarrow (\perp \rightarrow t_3)$	(modus ponens from 8, 9)
11	$\perp \rightarrow t_3$	(modus ponens from 5, 10)

12	t_4	(axiom 3)
13	$t_4 \rightarrow (\perp \rightarrow t_4)$	K
14	$\perp \rightarrow t_4$	(modus ponens from 12, 13)
15	$(\perp \rightarrow t_4) \rightarrow ((\perp \rightarrow t_3) \rightarrow (\perp \rightarrow t_5))$	S
16	$(\perp \rightarrow t_3) \rightarrow (\perp \rightarrow t_5)$	(modus ponens from 14, 15)
17	$\perp \rightarrow t_5$	(modus ponens from 11, 16).

Question 3

We want to show that $p \vdash (p \rightarrow \perp) \rightarrow \perp$. By the deduction theorem, it suffices to show that $\{p, p \rightarrow \perp\} \vdash \perp$. But this follows by *modus ponens*.

Question 4

We want to show that $\{p, q\} \vdash (p \rightarrow (q \rightarrow \perp)) \rightarrow \perp$.

(i) By the deduction theorem, it suffices to show that $\{p, q, p \rightarrow (q \rightarrow \perp)\} \vdash \perp$. But this follows by two applications of *modus ponens*.

(ii) By the completeness theorem, it suffices to consider a valuation ν with $\nu(p) = \nu(q) = 1$. Then $\nu(q \rightarrow \perp) = 0$, whence $\nu(p \rightarrow (q \rightarrow \perp)) = 0$, from which it follows that $\nu((p \rightarrow (q \rightarrow \perp)) \rightarrow \perp) = 1$.

(iii) By the proof of the deduction theorem, the following is a proof that $\{p, q\} \vdash p \wedge q$, where $x = (p \rightarrow (q \rightarrow \perp))$:

(1)	$x \rightarrow (x \rightarrow x)$	K
(2)	$x \rightarrow ((x \rightarrow x) \rightarrow x)$	K
(3)	$(x \rightarrow ((x \rightarrow x) \rightarrow x)) \rightarrow ((x \rightarrow (x \rightarrow x)) \rightarrow (x \rightarrow x))$	S
(4)	$(x \rightarrow (x \rightarrow x)) \rightarrow (x \rightarrow x)$	(modus ponens from 2, 3)
(5)	$x \rightarrow x$	(modus ponens from 1, 4)
(6)	p	(in S)
(7)	$p \rightarrow (x \rightarrow p)$	K
(8)	$x \rightarrow p$	(modus ponens from 6, 7)
(9)	q	(in S)
(10)	$q \rightarrow (x \rightarrow q)$	K
(11)	$x \rightarrow q$	(modus ponens from 9, 10)
(12)	$(x \rightarrow x) \rightarrow ((x \rightarrow p) \rightarrow (x \rightarrow (q \rightarrow \perp)))$	S
(13)	$(x \rightarrow p) \rightarrow (x \rightarrow (q \rightarrow \perp))$	(modus ponens from 5, 12)
(14)	$x \rightarrow (q \rightarrow \perp)$	(modus ponens from 8, 13)
(15)	$(x \rightarrow (q \rightarrow \perp)) \rightarrow ((x \rightarrow q) \rightarrow (x \rightarrow \perp))$	S
(16)	$(x \rightarrow q) \rightarrow (x \rightarrow \perp)$	(modus ponens from 14, 15)
(17)	$x \rightarrow \perp$	(modus ponens from 11, 16).

Now, from the premise $\neg p$, (or $p \rightarrow \perp$), together with a proof that $\perp \rightarrow q$ for arbitrary q , we conclude that $p \rightarrow q$ by the example in class.

(Qiaochu Yuan again)

Question 5

It suffices to set $q := \neg p$. Suppose there were a valuation ν such that $\nu((p \rightarrow \neg p) \rightarrow \neg(\neg p \rightarrow p)) = 0$. Then $\nu(p \rightarrow \neg p) = 1$ and $\nu(\neg(\neg p \rightarrow p)) = 0$, whence $\nu(\neg p \rightarrow p) = 1$. But if $\nu(p) = 1$, then the first condition is impossible, and if $\nu(p) = 0$, then the second condition is impossible; contradiction. So there exists no such valuation.

Question 6

Pay heed to the word ‘carefully’. (It would have been much better if Professor Leader had challenged you to *show how to count ...*). What he wants you to do is prove, by induction on n , that the set of formulæ of depth n is countable. He (and I, too) want you to do this by explicitly showing how to obtain an enumeration of the set of formulæ of depth $n + 1$ from an enumeration of the set of formulæ of depth n . That will give you an ω -sequence of enumerations which you can stitch together to obtain a wellordering of the union. The stitching together is done in the standard zigzag way that you use to enumerate $\mathbb{N} \times \mathbb{N}$. If you do it that way, then you have explicitly exhibited an enumeration of the language.

You will all of you want to prove by induction on n that the set of formulæ of depth n is countable, but you might feel inclined to appeal to the sirens you heard in Numbers and Sets who told you that a union of countably many countable set is countable, and to use that at each step in the induction, as well as in the final wrap-up stage. Even if that is true (and certainly there are people who believe it) it’s bad practice to appeal to it, beco’s (i) you don’t need it (as we have seen) and (ii) a proof that uses that principle contains less information than the constructive proof I have outlined above.

There are other cute ways of doing it. Here’s one of them. Structure your infinite set of primitive propositions as $\{p, pp, ppp, pppp \dots\}$. Your propositional language now has only *five* characters: ‘(’, ‘(’, ‘ \rightarrow ’, ‘ \perp ’ and ‘ p ’—rather than a countable infinity of them. Number these characters with the numbers 0 to 4. Now any number written in base 5 corresponds to a unique string from this alphabet. [For pedants: we don’t have to worry about leading zeroes beco’s no wff starts with a right parenthesis!] [Again—for pedants—the set we have shown to be countable is not the propositional language itself but rather a superset containing some ill-formed formulæ. However it is easy to recover a counting of the propositional language from this: after all, every infinite subset of \mathbb{N} can be effectively counted.]

That proof used the clever trick that made the alphabet finite, but you actually don’t need to do that. You can exploit unique factorisation of natural numbers to make every natural number encode a sequence of smaller natural numbers, namely the exponents of $2, 3, 5 \dots$ in its unique representation as a product of prime powers.

Question 7

This duplicates question 6 from sheet 1 Dr Russell’s set

Question 8

If we can deduce an expression ϕ from the first two axioms, where ϕ has occurrences of ‘ \perp ’, then we can also deduce the result of replacing in ϕ every occurrence of ‘ \perp ’ by some random propositional

letter not appearing anywhere in the proof. So if we could deduce $((p \rightarrow \perp) \rightarrow \perp) \rightarrow p$ we would be able to deduce $((p \rightarrow q) \rightarrow q) \rightarrow p$. At the risk of making a mountain out of a molehill I will, at this point, say that the set of things deducible from axioms 1 and 2 is an inductively defined set and supports an induction principle, and we can use this induction principle to show that everything in this set is a tautology: the two axioms are tautologies, and tautologousness is preserved by *modus ponens*. $((p \rightarrow q) \rightarrow q) \rightarrow p$ is not a tautology and therefore cannot be deduced from the first two axioms.

In earlier editions of this sheet there was a further question along these lines ... “if A is a tautology not containing ‘ \perp ’ must it be deducible from the first two axioms?”. This is a hard question. You might wish to pursue it. If you do, here is a slightly cuddlier version of it. “Find a tautology not containing ‘ \perp ’ which is not derivable from the first two axioms, and use structural induction on the inductively defined set of deductive consequences of the first two axioms to prove that underivability.” I have handouts on this with pretty pictures that it cost me blood to draw, so I’m hoping some of you will ask me about it.

Question 9

This duplicates q8 on Dr Russell’s sheet 1

Question 10

This duplicates question 11 on sheet 1 Dr Russell’s set.

Question 11

Let S be a set of propositions. We want to show that if $S \models t$, then S has a finite subset S' such that $S' \models t$. Suppose this is true whenever $t = \perp$. If $S \models t$, it follows that $S \cup \{\neg t\} \models \perp$, so there is a finite subset S' of $S \cup \{\neg t\}$ such that $S' \models \perp$. If S' does not contain $\neg t$, then it is a subset of S , so $S \models \perp$, hence $S \models t$. Otherwise, no valuation is equal to 1 on S' , so if a valuation ν is equal to 1 on $S' \setminus \{\neg t\}$ then $\nu(\neg t) = 0$, whence $\nu(t) = 1$, so $S' \models t$.

So it suffices to prove the claim when $t = \perp$. Let P be the set of primitive propositions. Since a valuation ν is determined by what it does on P , the set of valuations can be identified with the set $\{0, 1\}^P$. If $\{0, 1\}$ is given the discrete topology, then $\{0, 1\}^P$ is compact by Tikhonov’s theorem.

A proposition in L determines a function $f : \{0, 1\}^P \rightarrow \{0, 1\}$. Since the truth of a proposition can only depend on finitely many elements of P , any such function f has the property that the preimages of both $\{0\}$ and $\{1\}$ must be open, whence f is continuous.

Now let S be a set of propositions which determine a set of functions $f_s : \{0, 1\}^P \rightarrow \{0, 1\}$, $s \in S$. We are given that $S \models \perp$, whence there is no valuation which takes the value 1 on all of S . This is equivalent to the statement that the open sets $f_s^{-1}\{0\}$ form an open cover of $\{0, 1\}^P$ and, by compactness, this open cover has a finite subcover f_{s_1}, \dots, f_{s_n} . Then $\{s_1, \dots, s_n\} \models \perp$.

Question 12

Let $\{p_i : i \in \mathbb{N}\}$ be distinct primitive propositions. For $i \in \mathbb{N}$ define A_i to be $\bigwedge_{j \leq i} p_j$.

Clearly the A_i form an infinite chain.

An uncountable chain wrt deducibility? You must be joking.

Suppose we have uncountably many primitive propositions. Consider the symmetric group on the primitive propositions, and the orbits of its obvious action on compound propositions. Actually, on second thoughts, consider the subgroup consisting of those permutations of finite support (those that move only finitely many propositions). Why? Well, if the permutation σ moves a compound formula A to $\sigma(A)$ it does so only in virtue of a finite bit of σ so there will be a permutation of finite support that moves A to $\sigma(A)$. This will matter...

(Things belonging to the same orbit are said to be *alphabetic variants* [of each other; you may encounter this expression in other contexts] and the equivalence relation is sometimes called *α -equivalence*. In predicate calculus the existence of distinct-yet- α -equivalent formulæ is a pain, but it's one we get beco's we have variables.)

Now suppose *per impossibile* that we had an uncountable chain. Consider its intersections with the orbits. There are only countably many orbits. This is because each orbit corresponds to a “skeleton” of a formula—and there are only countably many skeletons.

The intersections of our putative chain with the orbits partitions it into countably many pieces. How big are the pieces? We want them to be so small that a union of countably many of them cannot be uncountable. Now you may know (and if you didn't you learnt it first here) that if AC fails badly enough then a countable set of pairs might have an uncountable sumset. So what we want to prove is that each piece is a singleton. That will do it.

Let A be a formula. Anything else in the orbit of A is $\sigma(A)$ for some permutation σ of finite support, and accordingly of order n , say, for some $n \in \mathbb{N}$. We claim that A and $\sigma(A)$ are either interdeducible or incomparable. Suppose not, and that $\vdash A \rightarrow \sigma(A)$. By composing our valuations (which are functions from primitive propositions to $\{T, F\}$) with the powers of σ we can see that we must also have $\vdash \sigma(A) \rightarrow \sigma^2(A)$, and $\vdash \sigma^2(A) \rightarrow \sigma^3(A)$ all the way up to $\vdash \sigma^{n-1}(A) \rightarrow \sigma^n(A) = A$. So any two comparable things in an orbit are interdeducible. So there are no chains even of length two, let alone uncountable chains!

■

Thanks to José Siqueira for compelling me to be clearer than i had been.

Actually here is another proof, due to Cong Chen. (Rather better than mine, in fact). This is not how he presents it, but the result of my doctoring. He does it in terms of probabilities, can you imagine! This is a *Logic* course for heaven's sake.

To each propositional formula with n distinct letters we can associate a rational number with denominator 2^n , namely the number of rows of its truth-table in which it comes out true divided by the number of rows in the truth-table. (OK, you can call it its probability if you insist). If $\phi \vdash \psi$ but not the other way round then the “probability” of ϕ must be less than the “probability” of ψ . Every valuation making ϕ true also makes ψ true. So the “probability” of ϕ is less-than-or-equal-to the “probability” of ψ . If the probabilities are the same then ϕ and ψ must be validated by the same valuations, and they ain't. This means that the map from the putative chain to the dyadic rationals is injective. And, as we all know, the set of dyadic rationals is countable, so the chain was countable.

So no uncountable chains. My guess is that you can get chains of any countable linear order type you like. And probably you can embed every countable partial order. No promises, mind.

Question 13*

Do not attempt this question. No, *really*.

Oh, all right: have a look at www.dpmms.cam.ac.uk/~tf/cam_only/rickard.pdf.

You see what i mean? Next time perhaps you'll believe me.

Some Old questions from last year which are still good for your soul

Question 11

This question duplicates question 12 on sheet 11 of Dr Russell's set.

Question 12

This first bit comes from Sean Moss, Senior Wrangler 2012 ... He has now bunked off to The Other Place. Boo! Hiss!!

For concreteness, we'll consider the length $l(\phi)$ of a formula to be the total number of primitive propositions (counted with multiplicity), and we won't worry about 0.

Main idea: for any formula ϕ , if v is any valuation then $v \models \phi$ or $v \models \neg\phi$.

Thus for any choice of pluses and minuses $\pm p_1, \dots, \pm p_m \vdash \phi$ or $\pm p_1, \dots, \pm p_m \vdash \neg\phi$, where the p_i are primitive propositions including all of those occurring in ϕ (and $\pm p$ means one of p or $\neg p$).

We first find a bound $g(n)$ on the length of a proof of $\pm p_1, \dots, \pm p_m \vdash \phi$ or $\neg\phi$.

We abbreviate $\pm p_1, \dots, \pm p_m$ as v (as in a valuation).

Claim We can take $g(n) = 2^{n+3} - 15$

Proof. If $n = 1$, then $\phi = p$ and $\phi = \neg p$ each have one-line proofs from $\pm p$, so $g(1) = 1$.

Suppose $\phi = (\psi \rightarrow \chi)$ and $v \vdash \phi$, $l(\phi) = n + 1$.

Then if $v \vdash \chi$, write down a $\leq g(n)$ -line proof of $v \vdash \chi$, followed by: $0 \quad \chi \rightarrow (\psi \rightarrow \chi)(K)$
 $\psi \rightarrow \chi(MP)$

If $v \vdash \neg\chi$, $\neg\psi$, then write down a $\leq g(n)$ -line proof of $\neg\psi$ followed by the 7-line proof of $\perp \rightarrow \chi$ and then the 6-line proof of $\psi \rightarrow \chi$.

Alternatively, if $v \vdash \neg(\psi \rightarrow \chi)$, then $v \vdash \psi, \neg\chi$.

Write down the two $\leq g(n)$ -line proofs of ψ and $\neg\chi$. Then

$$\psi, \neg\chi, (\psi \rightarrow \chi) \vdash \perp$$

in only five lines

ψ	(Hyp)
$\psi \rightarrow \chi$	(Hyp)
χ	(MP)
$\chi \rightarrow \perp$	(Hyp)
\perp	(MP)

By the proof of the deduction theorem, we can prove

$$\psi, \neg\chi \vdash \neg(\psi \rightarrow \chi)$$

in $3 \times 5 + 2 = 17$ lines.

Thus we can prove $v \vdash \neg(\psi \rightarrow \chi)$ in $2g(n) + 15$ lines (we save 2 by not repeating the hypotheses in the last 17 lines). Solving the recurrence gives us the stated bound. ■

Now we will use the fact that $T, p \vdash \phi$ and $T, \neg p \vdash \phi$ implies $T \vdash \phi$.

We can prove $\{p \rightarrow \phi, \neg p \rightarrow \phi, \neg\phi\} \vdash \perp$ in 10 lines:

$$\begin{aligned} 0 & \quad \phi \rightarrow \perp (Hyp.) \\ (\phi \rightarrow \perp) & \rightarrow (p \rightarrow (\phi \rightarrow \perp)) (K) \\ p & \rightarrow (\phi \rightarrow \perp) (MP) \\ (p \rightarrow (\phi \rightarrow \perp)) & \rightarrow ((p \rightarrow \phi) \rightarrow (p \rightarrow \perp)) (S) \\ (p \rightarrow \phi) & \rightarrow (p \rightarrow \perp) (MP) \\ p & \rightarrow \phi (Hyp.) \\ p & \rightarrow \perp (MP) \\ (p \rightarrow \perp) & \rightarrow \phi (Hyp.) \\ \phi & (MP) \\ \perp & (MP) \end{aligned}$$

By the deduction theorem there is a proof of $\{p \rightarrow \phi, \neg p \rightarrow \phi\} \vdash \neg\neg\phi$ in 32 lines.

Adding an instance of (T) and a (MP), we get a proof of $\{p \rightarrow \phi, \neg p \rightarrow \phi\} \vdash \phi$ in 34 lines.

Starting with N -line proofs of $\{\pm p_1, \dots, \pm p_{m-1}, p_m\} \vdash \phi$ and $\{\pm p_1, \dots, \pm p_{m-1}, \neg p_m\} \vdash \phi$ (where the \pm 's are fixed), use the deduction theorem to get $\leq (3N+2)$ -line proofs for $\{\pm p_1, \dots, \pm p_{m-1}\} \vdash p_m \rightarrow \phi, \neg p_m \rightarrow \phi$.

Add 32 lines to get to ϕ .

The process thus gives us a $(6N + 34)$ -line proof of

$$\{\pm p_1, \dots, \pm p_{m-1}\} \vdash \phi.$$

Since the number of primitive propositions in ϕ is bounded by its length, we need only iterate this a total of n times. Round up to $(6N + 35)$ for convenience and then the n^{th} iterate is $6^n(N + 42) - 7$.

Thus the final bound we achieve is

$$\begin{aligned} f(n) &= 6^n(2^{n+3} - 15) - 7 \\ &= 8 \cdot 12^n - 15 \cdot 6^n - 7 \\ &= O(12^n). \end{aligned}$$

Thank you very much, Dr Moss!

You might wonder whether this exponential bound is best possible. Curiously, this is an open question. I have to be careful how to state this, because I suspect that for this particular presentation of propositional logic it probably *is* best possible—and is known (tho' not to me) to be best possible. The open question is whether or not there is a proof system for propositional logic in which there is a polynomial bound on lengths of proofs.

This is related to the $P = NP$ question, or (more precisely) to the $NP = co\text{-}NP$ question. The set of falsifiable formulæ of propositional logic is in NP (guess a valuation, verify in linear time that it falsifies the candidate). This is because a set X of things is NP (“is an NP -set”) iff (by definition) you become a member of X in virtue of being related to something by an easily decidable relation; our example here is: you are refutable formula of propositional logic iff there is a valuation that refutes you. It’s actually *NP-complete*, which is to say it’s as bad a problem as an NP problem can be. (Every NP problem can be coded up—in polynomial time—as a question about satisfiability of a propositional formula). Now the set of tautologies is the complement of the set of falsifiable sentences, and thus is in $co\text{-}NP$. (A $co\text{-}NP$ set is one whose complement is an NP set). Now, if we could find a proof system for propositional logic in which every tautology had a proof of polynomial length, then the set of tautologies would be in NP : guess a proof, verify in time linear in the proof (polynomial in the candidate) that it is a proof of the candidate. So we would have a problem that is $co\text{-}NP$ and is NP -complete, so every $co\text{-}NP$ problem would be in NP whence $NP = co\text{-}NP$. This is an open problem . . . a **hard** open problem!

Sheet 2

Question 1

Write down subsets of \mathbb{R} of order types $\omega + \omega$, ω^2 and ω^3 in the inherited order.

One of my students came up with $\{1 - 1/n : n \in \mathbb{N}\} \cup \{10 - 1/n : n \in \mathbb{N}\}$. Why that rather than $\{1 - 1/n : n \in \mathbb{N}\} \cup \{2 - 1/n : n \in \mathbb{N}\}$, i wondered ...? His answer is the range of an order-preserving map from the ordinals below $\omega + \omega$ into \mathbb{R} . My preferred answer is the range of a *continuous* order-preserving map from the ordinals below $\omega + \omega$ into \mathbb{R} .

ω^2 is not that hard: $\{n - 1/m : n, m \in \mathbb{N}\}$, but ω^3 requires a bit of work. Fortunately most of you were up to it. The key observation is that, in each copy of ω , the gap between the m th and the $m + 1$ th point is $\frac{1}{m(m+1)}$ wide, so if you want to squeeze an extra copy of ω in there you do

$$\left\{n - \frac{1}{m} - \frac{1}{km(m+1)} : n, m, k \in \mathbb{N}\right\}$$

Test your comprehension by doing ω^4 in the same style.

Question 2

Does a picture serve for a proof for these equations? Depends partly on whether you are (i) trying to persuade yourself of the truth of the allegation (by gaining understanding) in which case it's probably all right, or (ii) trying to remove all doubt, in which case it might not be.

Question 3

Question 4

(Inductive and synthetic definitions)

This question goes to the heart of how to think of ordinals.

The correct way to prove that the two definitions are equivalent is to fix α and prove by induction on β that the two definitions agree on $\alpha \cdot \beta$.

Well it's obviously true for $\beta = 0$! (OK, it's trivial, but at least it's a start.)

Suppose $\beta = \gamma + 1$. Then the recursive definition tells us that $\alpha \cdot \beta = \alpha \cdot \gamma + \alpha$. But this is clearly the length of a wellorder (any wellorder) obtained by putting a wellorder of length α on the end of a wellorder of length $\beta \cdot \gamma$.

It's at the limit stage that we have to do some work. So suppose the inductive and synthetic definitions of $\alpha \cdot \gamma$ agree for all $\gamma < \beta$. Consider a wellorder that is of length $\alpha \cdot \beta$ according to the synthetic definition. Up to isomorphism we can think of it as the lexicographic product of $\langle A, <_A \rangle \times \langle B, <_B \rangle$ for two wellorderings $\langle A, <_A \rangle$ and $\langle B, <_B \rangle$ of lengths α and β . Now let γ be an ordinal below β . Every such ordinal is the order type (length) of a unique initial segment of $\langle B, <_B \rangle$; let us write this as $\langle B, <_B \rangle \upharpoonright \gamma$. Our lexicographic product $\langle A, <_A \rangle \times \langle B, <_B \rangle$ is now a colimit of all the $\langle A, <_A \rangle \times \langle B, <_B \rangle \upharpoonright \gamma$ for $\gamma < \beta$. Each $\langle A, <_A \rangle \times \langle B, <_B \rangle \upharpoonright \gamma$ is of length $\alpha \cdot \gamma$ —and that is according to *either* definition, by induction hypothesis. So the length of $\langle A, <_A \rangle \times \langle B, <_B \rangle$ must be the supremum of $\{\alpha \cdot \gamma : \gamma < \beta\}$ and this is the recursive definition of $\alpha \cdot \beta$.

Question 5

Ordinal multiplication is associative. The only sane way to prove this is by using the synthetic definition. In fact it is *always* best to prove facts about ordinals synthetically (wherever possible) rather than by induction. Let me say a bit about why this is so. Doing it by induction relies on the three order-types being ordinals, but that's not why it's true. It's true for *arbitrary* linear order types. You do it just by rearranging the brackets inside the two products; the fact that α , β and γ are ordinals is irrelevant and shouldn't be exploited!

If you want to do it by induction there are some things you should think about. For a start there are two kinds of induction you can do over the ordinals. There is structural induction, where you consider three cases: (i) $\alpha = 0$, (ii) α successor, and (iii) α limit. Then there is *wellfounded* induction where you prove that α is F as long as every smaller ordinal is F . These correspond to the two kinds of induction you can do over \mathbb{N} , and they are of course equivalent—just as those two kinds of induction over \mathbb{N} were. But in practice of course it's sometimes much easier to do it one way rather than the other.

Now suppose you are trying to prove that $\phi(\alpha, \beta)$ holds for all ordinals α and β . There are six ways you could do it.

- (i) Say: “let α and β be arbitrary”, reason about them, conclude the things you want
- (ii) You could fix α , and prove by induction on β that $(\forall \beta)(\phi(\alpha, \beta))$, where your induction hypothesis is $\phi(\alpha, \beta)$; then say “but α was arbitrary...”
- (iii) You could fix β , and prove by induction on α that $(\forall \alpha)(\phi(\alpha, \beta))$ where your induction hypothesis is $\phi(\alpha, \beta)$; then say “but β was arbitrary...”
- (iv) You could prove by induction on α that $(\forall \beta)(\phi(\alpha, \beta))$ where your induction hypothesis is $(\forall \beta)(\phi(\alpha, \beta))$;
- (v) You could prove by induction on β that $(\forall \alpha)(\phi(\alpha, \beta))$ where your induction hypothesis is $(\forall \alpha)(\phi(\alpha, \beta))$;
- (vi) You could perhaps do a wellfounded induction on the lexicographic product... infer $\phi(\alpha, \beta)$ from the assumption that $\phi(\alpha', \beta')$ holds for all pairs α', β' below α, β in the lexicographic product.

That's bad enough. The thing we are challenged to prove here has *three* variables in it. I don't want to think about how to do it by induction: life is too short.

Tho' i s'pose i ought to, really. I think the correct way to prove $(\forall \alpha \beta \gamma)(\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma)$ (if you are going to do it by induction) is by Universal Generalisation on ' α ' and ' β ' (“Let α and β be arbitrary”) and do an induction on γ . No promises, mind.

But, as I said at the outset, you should do it synthetically.

Question 6

Check these by thinking *synthetically*. It becomes very clear very quickly. Again it's worth pointing out that the equation $\alpha \cdot (\beta + \gamma)$ holds for all linear order types not just ordinals, so it can't be

right to try and prove it by induction.

Question 7

“Ordinal subtraction is defined synthetically by taking $\alpha - \beta$ to be the order-type of the set-difference $\alpha \setminus \beta$ (in particular, $\alpha - \beta = 0$ whenever $\alpha \leq \beta$). Prove the following identities:

$$(\alpha + \beta) - \alpha = \beta \quad ; \quad \alpha - (\beta + \gamma) = (\alpha - \beta) - \gamma \quad ; \quad \alpha \cdot (\beta - \gamma) = \alpha \cdot \beta - \alpha \cdot \gamma \quad .$$

Show also that for any ordinal α there are only finitely many ordinals of the form $\alpha - \beta$. [Hint: consider the order-type of the set $\{\alpha - \beta : \beta \in \mathbf{On}\}$.]”

Discussion

Let’s think about how to prove that $(\beta + \alpha) - \beta = \alpha$. You take a well order of length α and stick it on the end of a well order of length β . Then you remove from the bottom end a well order of length β and what is left (the thing that is going to be of length $(\beta + \alpha) - \beta$) is obviously the thing of length α that you stuck on the end in the first place.

For this operation of ordinal subtraction to be well defined we need to be confident that when we subtract β (by chopping off an initial segment of length β) then there is only one initial segment of length β to chop off. If there is more than one, then there might be more than one answer to the question “What is $(\beta + \alpha) - \beta$?” This is where we have to exploit the fact that we are dealing with wellorderings. After all if we try to subtract ω^* from ω^* (ω^* is ω “upside-down”, the order type of the negative integers) we can get any natural number. The uniqueness we want can be had because we are dealing with ordinals not arbitrary linear order types.

Let’s prove it.

If $\alpha \geq \beta$ then any well order $\langle A, < \rangle$ of length α has a initial segment of length β . We need this initial segment to be unique. Suppose $\langle B, < \rangle$ is a well order of length β . It is isomorphic to an initial segment of $\langle A, < \rangle$, and if it were isomorphic to more than one initial segment of $\langle A, < \rangle$ then the isomorphism between the two would give rise to an subset of A with no least element. (Think of the trajectory under the isomorphism of an element in the symmetric difference of the two initial segments). The length of the terminal segment is $\alpha - \beta$. The uniqueness of the initial segment ensures that the terminal segment is unique, so this operation is well-defined.

[the rest of this discussion concerns the uniqueness of subtraction, something that was deleted from the current version of this question. I’m leaving it in coz it’s good for your soul.]

Why, for each α , are there only finitely many ordinals of the form $\alpha - \beta$? Why is this so plausible? Well, suppose that, for some α , there were infinitely many ordinals of the form $\alpha - \beta$. Then there would be infinitely many β with all the $\alpha - \beta$ distinct. Lots of β might give the same $\alpha - \beta$ so just pick the least. But then these finitely many β must form an increasing sequence. Reflect now that $\beta \mapsto \alpha - \beta$ is *decreasing* is *antimonotonic* so we would get an infinite *decreasing* sequence of values of $\beta \mapsto \alpha - \beta$.

I have struggled to find the cutest proof of this fact, but in the final analysis i decided it’s best to do it by induction on ordinals.

Let α be the smallest ordinal s.t. there are infinitely many ordinals of the form $\alpha - \beta$, and let $\beta < \alpha$ be the first ordinal s.t $\alpha - \beta \neq \alpha$. Observe now that every ordinal of the form $\alpha - \delta$ is of the form $\alpha - (\beta + \gamma)$, which is to say of the form $(\alpha - \beta) - \gamma$. Now there are infinitely many ordinals of the form $\alpha - \delta$ but only finitely many of the form $(\alpha - \beta) - \gamma$.

Intermezzo

Some thoughts and advice is in order on this first crop of questions on ordinals and order types. It's a racing certainty that there will be a question about ordinals in your Part II exams. I am not in favour of mark-grubbing but it seems pointless to turn down a free α . You will be asked questions about equations and inequations, and invited to prove the true ones and find counterexamples to those that are false. Some of the true ones (like distributivity on the right of \times over $+$, and associativity of \times and $+$) work for arbitrary linear order types and therefore can be proved by hand and you don't need induction. Some of them work only for ordinals and then you need to exploit the fact that you are dealing with ordinals. $\alpha + 1 > \alpha$ is true for ordinals but not for arbitrary linear order types (think of ω^*) *so you have to exploit somehow the fact that α is an ordinal*. Exploiting the fact that the characters in your play are ordinals doesn't necessarily mean you have to be doing an *induction*—see the discussion of uniqueness of subtraction above.

Question 8

You want three tosets none of which embeds in either of the others? Piece of cake. The rationals, the countable ordinals and the countable ordinals turned upside-down. This question is from Dr Russell, and I don't know what examples he had in mind. Prof Leader wants to make lots of things of order type ω and ω^* and add up finitely many of them in annoying ways. That's probably more to your taste—and it certainly works. I think with a little work you can show—just using lots of copies of \mathbb{N} and \mathbb{N} upside-down (the negative integers)—that you can get finite antichains as wide as you like. Here's how to get an antichain of width 2^n . Take all your n -bit words, and in each replace the 0s by ω and the 1s by ω^* , and concatenate them. (Thus, when $n = 2$, you get: $\omega + \omega$, $\omega + \omega^*$, $\omega^* + \omega$ and $\omega^* + \omega^*$). Can you get infinite antichains? Think about what happens if you have things like this made from ω pieces strung together. You don't get an infinite antichain! Yes you *can* get infinite antichains, but in every infinite antichain there must be at least one total ordering of an uncountable set. This is corollary of a beautiful theorem of the late and much lamented Richard Laver. Some years I set a Part III essay on it. If you want to have a look at it (and it is very nice) then point your search engine at *Laver's proof of the Fraïssé conjecture*. This has connections with Question 14 on this sheet.

Question 9

If α is a countable nonzero limit ordinal, it is the order type of a wellordering $<_\alpha$ of \mathbb{N} . You now have *two* wellorderings of \mathbb{N} . You construct an increasing ω -sequence of naturals by “picking winners” (Prof. Leader's expression). Set a_0 , the first member of the sequence, to be 0; thereafter a_{n+1} is to be the $<_{\mathbb{N}}$ -least natural that is $>_\alpha a_n$. Now set α_i to be the length of the initial segment of $\langle \mathbb{N}, <_\alpha \rangle$ bounded by a_i .

Actually Michael Savery has a rather cute formulation of this. He says a natural number n is “tall” iff $(\forall m <_{\mathbb{N}} n)(m <_\alpha n)$, and he gets his sequence of α_i from the tall naturals.

For the moment i'm going to leave it to you to verify that we never run out of naturals, and that the sequence $\langle a_i : i \in \mathbb{N} \rangle$ is unbounded in $<_\alpha$. The sequence of ordinals that you have obtained is a **fundamental sequence for α** . This shows that every countable limit ordinal has cofinality ω .

(Actually it shows slightly more than that: notice that we did not exploit the assumption that α is an *ordinal*. All we used was that it was the order type of a countable total ordering with no

last element.)

Essentially the same proof (perhaps slightly neater) starts with the reflection (going back to Cantor) that each ordinal α is the ordertype of the set (which i think Professor Leader notates ' I_α ') of the ordinals below α in their natural order. If α is a countable ordinal then I_α is a countable set, so you exploit a counting of it (a bijection with \mathbb{N}) in the same way. That way you get the fundamental sequence directly. But it's the same proof really.

The interesting fact about this question is that you cannot compute the ω -sequence-of-smaller-ordinals-whose-supremum-is- α from α itself; you can only compute it from, so to speak, a *manifestation* of α , a wellordering of \mathbb{N} of length α . One is thrown off the scent by the fact that in some cases (in fact in all cases known to you so far) it's perfectly obvious what the ω -sequence should be: for ω^ω it's $\langle \omega^n : n < \omega \rangle$, for ϵ_0 it's $\omega, \omega^\omega, \omega^{\omega^\omega} \dots$. The problem is that there is no distinguished counting of I_α . There are countings all right (lots of them)¹ but no *distinguished* countings.

In the construction above, the particular ω -sequence you end up with will depend on your choice of $<_\alpha$. How many such $<_\alpha$ are there? (The answer to this riddle is not important, but I want you to be able to compute it)

Observe that Set Theory is no help here. It's true that each countable ordinal has a canonical representative—in the form of the corresponding von Neumann ordinal—but this is no help, beco's these von Neumann ordinals do not come equipped with canonical bijections with \mathbb{N} !

Finally you might like to check your comprehension by proving analogously that every limit ordinal between ω_1 and ω_2 is a limit of either an ω -sequence or an ω_1 sequence of smaller ordinals.

Question 10

Question 9

(Tripos II 93206). For each countable ordinal α , show that there is a subset of \mathbb{R} which is well-ordered (in the usual ordering) and has order-type α . Is there a well-ordered subset of \mathbb{R} (again, in the usual ordering) of order-type ω_1 ?

It works not just for countable ordinals, but any countable order type whatever!

Take any total order of \mathbb{N} . We will define an injection into \mathbb{Q} by recursion on the naturals. Send each natural number as it pops up to, well, the first positive integer if it is to the *right* of stuff already allocated, or the first negative integer if it is to the *left* of stuff already allocated. If it is between two things already allocated send it to the arithmetic mean of the things its immediate upper and lower neighbours were sent to.

That's the correct way to do it. There is a wrong way to do it, which most people pounce on, and that is to try to do it by induction on countable ordinals. It works, but you have to use countable choice to pick fundamental sequences for all limit ordinals. I shall spare you the details, since you may well have worked them out for yourself. However i will spell out an amusing detail.

Let us suppose that we have—by the above ruse, using countable choice—obtained a family $\langle f_\alpha : \alpha < \omega_1 \rangle$ where f_α injects the ordinals below α into \mathbb{R} in an order-preserving way. Fix a countable ordinal ζ and consider the ω_1 -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

¹How many?

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!

REMARK 2. For each countable ordinal γ , the sequence $\langle f_\gamma(\zeta) : \gamma > \zeta \rangle$ is not monotone nonincreasing.

Proof: Suppose that

$$(\forall \gamma < \gamma' < \omega_1)(\forall \zeta < \omega_1)(f_\gamma(\zeta) \geq f_{\gamma'}(\zeta)). \quad (2.1)$$

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 $\alpha < \beta \rightarrow f_\infty(\alpha) < f_\infty(\beta)$. (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.² ■

But that appealed to the second part of the question, which i had better now prove.

For the second part ("can you do the same for ω_1 ?") ...

There can be no subset of \mathbb{R} that is of order-type ω_1 in the inherited order. Suppose S were such a set. Observe that to the right of every element of S is an open interval disjoint from S . That is to say \mathbb{R} is naturally partitioned into half-open intervals, and this partition is in 1-1 correspondence with S , each member of S being paired with the half-open interval of which it is the left endpoint. This partition can be injected into \mathbb{Q} by sending each piece to the first rational in it. So S was countable after all.

I have noticed that a surprising number of you use arguments involving countable choice.

One such argument says that, if there were a set X of reals of order-type ω_1 in the inherited order then each of the intersections $X \cap (n, n+1]$ would be countable, meaning that X is a union of countably many countable sets and is therefore countable, contradicting the assumption that it is of length ω_1 and therefore of size \aleph_1 .

Using AC is bad practice even if AC is true. You don't want to use just any true fact that happens to be lying around: "God exists, so there is no order-preserving map from the second number class into the reals" doesn't quite cut it.

Some of you even managed to muck up the proof of two paragraphs above. OK, you send each countable ordinal to the open interval in \mathbb{R} as above. You then say: each interval contains a rational—which indeed it does—and then shut up shop and go home. That's not really good enough. The contradiction comes from having a function from a set of size \aleph_1 (the set of countable ordinals "the second number class") into a set of size \aleph_0 (the rationals). You can't stop until you have done it. You have to actually pick a rational from each of these intervals, so that you can send the countable ordinal in question to that rational. Which rational? With many of you it cost blood and threats of the rack to get you to say that the rationals have an ordering of length ω so you

²I suspect that the the sequence $\langle f_\gamma(\zeta) : \gamma > \zeta \rangle$ of values given to ζ describe a nonmeasurable set. I have seen no proof of this, tho'. We needed AC to build it so it might well be nonmeasurable.

pick, from each interval, the first rational in that interval in the sense of that wellordering. Even after I had spelled this out, a lot of you clearly just thought I was barmy. Well, I'm not: what I was trying to get you to do was come up with a proof, not a nondeterministic add-warm-water-and-stir pseudoproof. That's Logic for you!

More temperately [calm down and breathe deeply, tf] what is going on here is that we want to prove that, were there *per impossibile* an object of the conjectured kind (to wit, an order-preserving injection from the second number class into the reals) then there would be an object of a kind we know there cannot be, namely an injection of an uncountable set into a countable one. The proof must describe such a construction of an object of the second kind from an object of the first kind. One should never be *completely* satisfied with a nondeterministic construction if a deterministic construction is available.

If you want to think more about this have a look at chapter 2 (pp 20 ff) of www.dpmms.cam.ac.uk/~tf/fundamentalsequence.pdf

One of the things that this shows is that the quasiorder of linear order types (quasiordered by injective homomorphism) is not complete, or anything remotely like it: ω_1 and η (the order type of \mathbb{Q}) are distinct upper bounds for the second number class. ω_1 is a *minimal* upper bound but it is not the *minimum* upper bound, co's it ain't less than \mathfrak{c} . \mathfrak{c} (the order type of the reals) is an upper bound, but it is not a *minimal* upper bound; there is an infinite strictly descending sequence of upper bounds for the second number class all below \mathfrak{c} . (This is a theorem of Sierpinski, using a grubby diagonal argument powered by a wellordering of \mathbb{R} . I used to lecture it in my Part III lectures on WQO theory. It also shows its face in an Impossible Imre Question (question 14 sheet 2, 2015). Indeed this question remains undead, since it is Q14 on This Very Sheet. The IIQ is "Suppose $\langle X, \leq_X \rangle$ is a total order with no non-identity injective homomorphism into itself. Must X be finite?")

Actually it's even worse than that: the quasiorder of linear order types isn't even a poset, beco's antisymmetry fails! (Consider $(0, 1)$ and $[0, 1]$.)

Question 11

Recall that a normal function is a function that is strictly increasing and continuous at limits.

To prove Cantor's Normal Form theorem we will need to make frequent use of the following important triviality. On is the collection of all ordinals. Don't worry at this stage about whether it's a class or a set.

REMARK 3. *If $f : On \rightarrow On$ 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 $\sup\{\alpha : f(\alpha) \leq \beta\}$. By continuity of f

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

which, by continuity of f , is

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

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 3 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

$$\gamma \mapsto \alpha + \gamma; \quad \gamma \mapsto \alpha \cdot \gamma; \quad \gamma \mapsto \alpha^\gamma$$

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 3 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

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

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 3 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

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

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:

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

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

So we have proved this:

THEOREM 2. *For all α and β there are $\gamma_0 > \dots > \gamma_n$ and $\theta_0 \dots \theta_n$ with $\theta_i < \alpha$ for each i , such that*

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

■

We can also prove it by using only the first part of this proof, by extracting the largest power of α that is less than β and subtracting it—thereby obtaining something smaller—and appealing to induction.

In particular, if $\alpha = \omega$ all the θ_i are finite. Since every finite ordinal is a sum $1 + 1 + 1 + \dots$ 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 3 to some $\alpha < \epsilon_0$ the output of this process must be an expression containing ordinals below α .

I think you can also do it by wellfounded induction over $<_{On}$. (NOT by the two-flavoured induction that considers successors and limits).

Suppose every ordinal $< \alpha$ has a CNF. Then either α is a power of ω (in which case we have a CNF immediately) or it isn’t ... in which case it’s a sum of two smaller ordinals, and again we get a CNF. Why is it a sum of two small ordinals? beco’s there is a maximal β s.t. $\omega^\beta \leq \alpha$!

In this treatment one then has to prove that the CNF one obtains is unique. (One should really do it anyway, but in the other treatment it’s sort-of obvious that it’s unique). I might try to find something enlightening to say about this later.

Question 12

This is hard. If it were me putting this sheet together i’d star it, but Prof Leader is Hungarian. Say no more. If there is a key steer on this then it’s the thought that if you want $\alpha * \beta$ to be always defined then you need a principle that says that every normal function not only has a fixed point but has arbitrarily late fixed points. It’s not difficult to persuade yourself that this is true, but turning this intuition into a proof is a surprisingly tricky affair, and results in a delicate expository problem for the lecturer. Proving this principle properly requires an appeal to the axiom scheme of replacement. This is beco’s the fixed points that you want are obtained as the sup of particular sequences of ordinals. The problem is: how do you know that those sequences actually exist? (And if the sequences exist how do you know they have sups? They do, but why?) Another thing to think about is: how do you know that all these buggers are countable? Are they, indeed?

Some students will be happy with a fairly informal proof that just ignores these issues in favour of just pressing on with the *idea* whereas more thoughtful (or more anxious) students may worry about the fact that we seem to need set theoretical principles that we haven’t seen yet and won’t see until the last quarter of the course. When i lecture this stuff i leave Q12 material until after we’ve seen some set theory, but perhaps i’m being over cautious ... plenty of you seem to cope quite well. In any case there is no one right way of doing it.

A word on motivation. Q 12 is the result of a line of thought arising from the fact that Cantor Normal Form isn’t always informative. It “crashes” at ϵ_0 , as you have seen. Of course you can

restart at ϵ_0 (using ϵ_0 as the base for your exponentiation instead of ω) but then you crash at ϵ_1 . And so on. What you want is a system of notation that doesn't have to have its tyres and oil changed every time you reach a fixed point for $\alpha \mapsto \omega^\alpha$. That's what the $*$ system of notation does.

And here is the rather scary thought. How many ordinals can you notate using just the symbols '0', '*', and '+' and ' α th fixed point of f '? Obviously only countably many...and there are uncountably many countable ordinals (how many??). So at some point this system of notation will crash too. In fact *any* system of notation for countable ordinals will crash sooner or later. Look back at Q11 and your proof, by induction on countable ordinals, that for every countable α there is a set of reals of that order type. At limit α you needed a fundamental sequence for α . But you get fundamental sequences for ordinals from systems of ordinal notations (that reach that ordinal). But no system of ordinal notation covers all countable ordinals, so you cannot *uniformly* assign fundamental sequences to countable ordinals. That's why you need AC!

Look at www.dpmms.cam.ac.uk/~tf/fundamentalsequence.pdf. You could also look up the *Veblen hierarchy*.

Question 13

Can there be an injective function f from countable limit ordinals to countable ordinals which is “pressing-down”— $f(\alpha) < \alpha$ always?

You might like to try to exhibit such a function. What's the obvious way to try? Greedily. Send ω to 0, then $\omega \cdot 2$ to 1, and so on up to ω^2 to ω . The process crashed at (i think) ω^ω . But if we use up the evens before the odds, then we don't use up all the finite ordinals until ω^3 (i think). One is left with the impression that if one uses up the naturals in a very funny order then one can postpone the crash out to a countable ordinal as big as one likes, but any algorithm of that flavour will crash eventually.

(Joel Tay's answer.)

Suppose f : countable limit ordinals to countable ordinals is pressing down and 1-1. This f organises the second number class³ into a family of disjoint descending ω^* sequences of limit ordinals ending in a successor ordinal; each member of the family is $\{f^{-n}(\alpha) : n \in \mathbb{N}\}$ for some successor α . Now join the family up by sending each successor ordinal to the last limit ordinal below it. This organises the whole of the second number class into a gigantic tree which is countably branching and of height ω . A wee bit of AC_ω is enough to secure a contradiction...specifically you prove by induction on the levels (and there are ω of them) that each level is countable

It is not hard to show that nevertheless such an f can always be found for any initial segment of the second number class. Think of a countable ordinal α . Well order \mathbb{N} in order type α . Send the limit ordinal β to the β th natural in this funny order. This leaves some room, so you can repeat the process, but if you iterate ω times you reach a fixed point.

Here is Professor Leader's proof.

Let f be an injective function defined on countable limit ordinals, and “pressing-down” $f(\alpha) < \alpha$. Set β_0 to be ω . Thereafter set $\beta_{n+1} := \sup\{\alpha : f(\alpha) < \beta_n\}$. (Observe that $f(\omega) < \omega$). f is injective, so the set of which β_n is the sup is countable, so (using AC_ω which tells us that any sup

³Cantor's name for the set of countable ordinals

of countably many countable ordinals is countable—miniexercise: why?), β_n is countable. Then consider what f must do to $\beta_\omega := \sup\{\beta_n : n \in \mathbb{N}\}$. Need i say more.

Actually i *do* need to say a bit more. As Catherine Willis of Pembroke has been astute and unkind enuff to point out, there is a problem with the definition of β_n given above, in that n might be such that, for all α , $f(\alpha) \leq \beta_n$ might imply $\alpha \leq \beta_n$, so we have to cast our net out a little further. I might try to sort out this glitch. On the other hand i might leave it to you.

Observe that this proof works even if f is allowed to be countable-to-one. We need AC_ω , but we needed it all along anyway.

This fact is a favourite fact of Prof. Leader's; it has something of the flavour of “ordinals are wellfounded” but in spades. And very useful it is too—look up *Fodor's* theorem and *Neumer's* theorem. This question contains the germ of the proofs of those two results.

I think (and you might like to prove this) that if α is any countable limit ordinal at all then, as long as you have assigned fundamental sequences to every limit ordinal below α one can spin f out to last at least for the ordinals below α .

Question 14⁺

This is a *lovely* question. When it was sprung on me the first time Prof Leader lectured this course I didn't know the answer, and it took me a long time to work it out. Once you know what the answer is, it's not *that* hard to do it, but how can you tell what this answer is? This uncertainty gives you the flavour of research mathematics.

If you are reading this then you probably didn't manage to work it out either, so you may be in the market for a hint!

Hint: Precisely how many order-preserving injections are there $\mathbb{R} \rightarrow \mathbb{R}$?

If you want another hint you should bear in mind that any set that can be wellordered at all can be wellordered in such a way that all its proper initial segments are smaller (have lesser cardinal) than the whole set. (Can you prove this fact to our shared satisfaction?)

I have a discussion answer to this question, but i have removed the link to it in response to Prof Leader's entreaties. However i am very happy to show it to you if you are seriously interested. There is some very interesting mathematics hidden behind this question.

Sharkovsky's ordering

I get $\omega^2 + \omega^*$

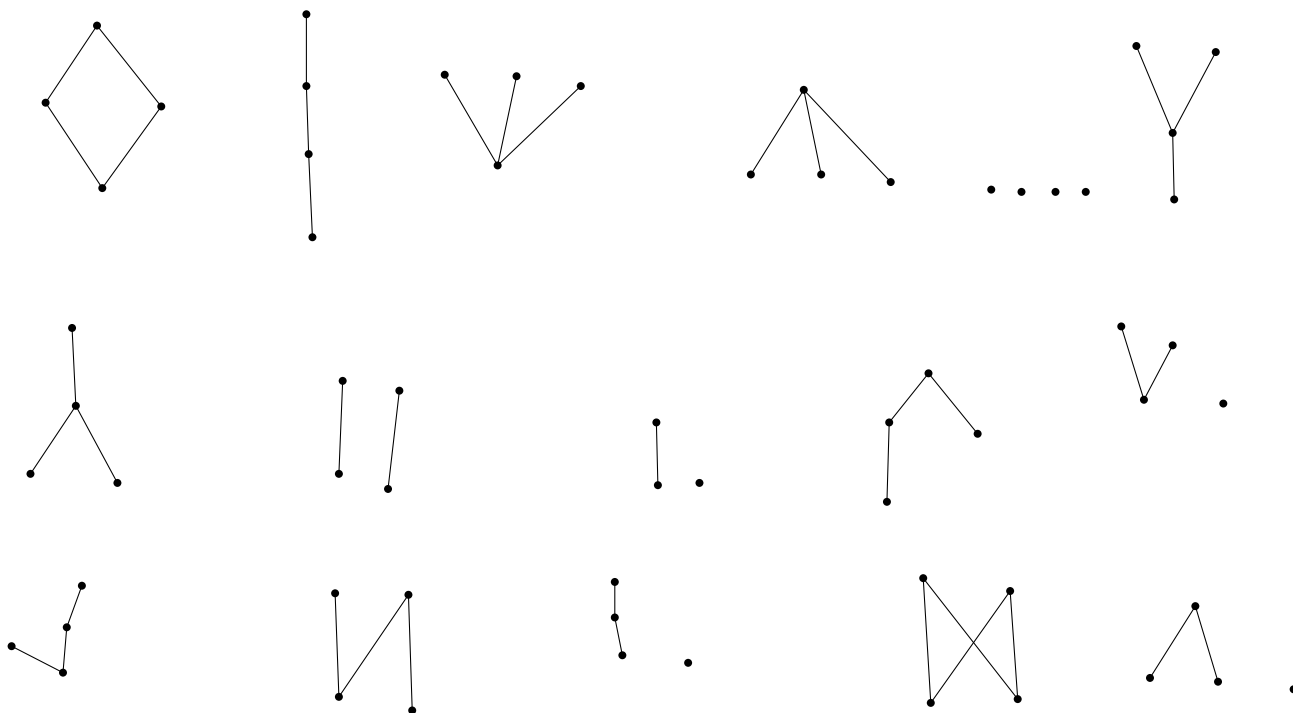
Sheet 3

Question 1

Is there a systematic way of generating them? No easy way, that's for sure. (The set of Hasse diagrams for posets with four elements is a quotient of the set of posets with four elements and in general the cardinalities of quotients are hard to compute, so you really don't know how many you are looking for, and you can't easily know when you have found them all.) I find myself wondering how many isomorphism types of partial orders there are on a set of n elements. Not *exactly* of course, for i see no prospect for an exact formula, but it would be nice to know whether or not there is an exponential lower bound, or a polynomial upper bound. There are $2^{\binom{n}{2}}$ reflexive relations on n elements. How many of them are transitive? My guess is: exponentially many, but i have no exact figure. Let me see... How many transitive relations on $n + 1$ things does a given transitive relation on n things extend to? There are $2n$ places where one might put in an edge. The only constraints arise as follows.

- (i) Suppose there is no edge from x to y , where x and y are of the original party of n . Then we cannot have an new edge from x to a (a is our new chap) as well as a new edge from a to y .
- (ii) If there is an edge from x to y and we add a new edge from y to a then we have to add a new edge from x to a .

Anyway, there are 16 isomorphism classes of posets on 4 elements, 2 of which are complete (the two with both a top and a bottom element, co's the empty subset has to have a sup!)



One of my students distinguished



... which are two embeddings of the same poset into the plane. This makes the same interesting point that my Pittsburgh colleague Ken Manders likes to make. When you formalise (= represent something concretely, or *concretise*) you add extra structure and this structure may be spurious. However I don't think this was the point that the question setters were trying to make... apparently the *real* reason for this question is that you weren't taught about Hasse diagrams in 1a. What is the world coming to??

There is a general question here: *How do i know when i've got them all?* This particular instance (before us) of this *general* question isn't so hard that we are prompted to think much about the general question, but a bit of thought won't go amiss. The answer of course is that you have to find a fairly robust way of thinking of these things as mathematical objects and then find a way of classifying them. In this case the obvious thing to do is to identify them with their Hasse diagrams and then classify them—perhaps—in terms of the number of edges they have. But the question still lurks in the shadows: “How can i give a *mathematical* proof that i have got all of them?”.

Back in the 1930s there was an American crime writer called *John Dickson Carr*, who specialised in locked-room murders. Some people say that his “The Hollow Man” is the best locked room murder of all time. In one of his novels (i forget which) his detective delivers himself of a long disquisition in the form of a classification of all locked-room murders. There is a small finite set of them apparently. I don't know how he could be sure, and i keep hoping to find a new one. It's the same with tragedies. Some Russian structuralist in the 1920's has a classification of them—again a small finite number.

A live version of this problem was the problem solved—within my lifetime—of the classification of all finite simple groups. How did the Monster crew know they'd got all of them? There is an answer to this, but i don't know it.

Question 2

Which of the following posets are complete?

(i) **The set of finite and cofinite subsets of \mathbb{N} , ordered by inclusion.**

It's not a complete poset, since the set $\{\{1\}, \{3\}, \{5\}, \dots\}$ does not have a supremum. That example also shows that it is not chain-complete.

(ii) **The set of independent subsets of a given vector space.**

The two elements $\{(1, 0), (0, 1)\}$ and $\{(1, 0), (1, 1)\}$ do not have a supremum, since any upper bound must include their union, and that is not linearly independent. However the collection of independent subsets of a vector space is of course *chain*-complete.

(iii) The set of subspaces of a vector space, ordered by set-inclusion.

This poset is complete. The supremum of any subset is the subspace spanned by the union of its elements.

(Observe that the *sup* and *inf* of this complete poset do not distribute. This is beco's **inf** is “honest” [it's just \bigcap] but **sup** is not: it's sometimes bigger than \bigcup .)

Some of you, I notice, want to pick a basis for each subspace and then take the union of the bases. This is unnecessary, and indeed undesirable. The point is not that it uses the axiom of choice—tho' it does—which is never a good idea if you can avoid it; the point is that it's also a violation of the vector-space rule that you should always prefer basis-independent proofs wherever they are available.

Question 3

The nicest and most natural example of an order-reversing map with no fixed point is complementation in a boolean algebra.

For the second part, if f is an order-reversing function from a complete poset into itself then f^2 is order preserving and has a fixed point.

Why on earth would you be looking for an order-reversing function to have a fixed point? More often than you might think. (And I don't just mean trivial cases like $1/2$ is a fixed point for the order-reversing function $x \mapsto (1 - x)$.) If you think a *species* in Biology is defined in terms of “can mate to produce viable offspring” you rapidly discover a characterisation in terms of fixed points for an order-reversing function. Have a look, too at this old tripos question (It was 2002:B2:11b).

1. State Zorn's lemma.
2. Let U be an arbitrary set and $\mathcal{P}(U)$ be the power set of U . For X a subset of $\mathcal{P}(U)$, the **dual** X^\vee of X is the set $\{y \subseteq U : (\forall x \in X)(y \cap x \neq \emptyset)\}$.
3. Is the function $X \mapsto X^\vee$ monotone? Comment.
4. By considering the poset of those subsets of $\mathcal{P}(X)$ that are subsets of their duals, or otherwise, show that there are $X = X^\vee$.
5. What can you say about the fixed points of $X \mapsto X^\vee$ on the assumption that U is finite?

Question 4

“Give the set of partial orders on S the containment partial order as subsets of $S \times S$. The resulting partial order is chain-complete, since the union of a nested sequence of partial orders is still a partial order. To see this, let \leq_n be a nested sequence of partial orders. The union partial order \leq is clearly reflexive. It is antisymmetric because $x \leq y$ and $y \leq x$ if and only if $x \leq_n y$ and $y \leq_m x$ for some m, n , and then it follows that $x =_{\max(m,n)} y$, whence $x = y$. Similarly, it is reflexive because $x \leq y$ and $y \leq z$ if and only if $x \leq_n y$ and $y \leq_m z$ for some m, n , and then it follows that $x \leq_{\max(m,n)} z$, whence $x \leq z$.”

That preceding paragraph was written by an earlier supervisee of mine, workname QY. It's fine, of course, but there is one point worth making He's trying to show that the poset is chain-complete (which it is). But we have no authority to assume that all chains are ω -sequences. You might have chains indexed by the rationals, or the countable ordinals, or by God-knows what.

Fortunately when you are trying to show that the union of a chain of partial ordering is another partial ordering you don't need any special conditions on the chain. It's true for any chain.

By Zorn's lemma, it follows that there exists a maximal partial order \leq' containing any given partial order \leq on S . For any $x, y \in S$, if x, y are incomparable then \leq' is not maximal since we can take the transitive closure of \leq' together with the relation $x \leq y$ to obtain a partial order strictly containing \leq' , so x, y are comparable and \leq' is a total order.

You can also do it by considering the poset of **total** orders of subsets of S that are compatible with the given partial ordering.

Question 5

Zorn's Lemma for countable posets.

You use the enumeration to ensure that the process of trying to reach a maximal element will succeed in finitely many steps.

Let $\langle X, \leq_X \rangle$ be a countable chain-complete poset. Enumerate X as $\langle x_i : i \in \mathbb{N} \rangle$. Build a \leq_X -chain the subscripts of whose elements form an $\leq_{\mathbb{N}}$ -increasing sequence. First one is x_0 , thereafter if the x -in-hand is maximal, then **HALT**; **else** plonk on the end that x which is \geq_X the x -in-hand which has $\leq_{\mathbb{N}}$ -minimal subscripts. If this doesn't **HALT** in finitely many steps the resulting chain has an upper bound and one obtains a contradiction by enquiring about the subscript on the upper bound.

I think the point of this question is to prepare you for a proof of ZL from AC. You want to show that a chain-complete poset $\langle X, \leq_X \rangle$ has a maximal element? Brutally wellorder X and use the technique of question 5.

Question 6

\Leftarrow : AC implies Zorn's lemma, which we then apply to the chain-complete poset of partial bijections between two given sets.

\Rightarrow : Let X be a set. By Hartogs' lemma, there exists a well-ordered set α with no injection $\alpha \rightarrow X$. It follows that there exists an injection $X \rightarrow \alpha$ which identifies X with a subset of α , which is itself well-ordered; thus X can be well-ordered.

Let S_i be a collection of sets indexed by an index set I , and choose a well-ordering on $\bigcup S_i$. For every i , let $f(i)$ be the least element of S_i relative to this well-ordering. Then $f(i)$ is a choice function.

Question 7

Zorn's lemon. Alternative answer: The Wellordering Pineapple.

Another suggestion (from Donald Hobson) is the Banach-Tarski paradox. But the Banach-Tarski paradox is strictly weaker than AC, so that can't be right.

The subtext to all this (as one of you were good enough to point out) is that it is *cowardly* to use Zorn's lemma. But perhaps *lazy* would be better.

Question 8

(i): Fields of Characteristic 2

The language has $\Omega = \{+, \times, 0, 1\}$ with arities 2, 2, 0, 0 and $\Pi = \emptyset$. The theory can be described by the following axioms:

$$\begin{aligned} &(\forall x)(\forall y)(x + y = y + x) \\ &(\forall x)(\forall y)(\forall z)((x + y) + z = x + (y + z)) \\ &(\forall x)(x + 0 = x) \\ &(\forall x)(\forall y)(x \times y = y \times x) \\ &(\forall x)(\forall y)(\forall z)((x \times y) \times z = x \times (y \times z)) \\ &(\forall x)(x \times 1 = x) \\ &(\forall x)(x \neq 0 \rightarrow (\exists y)(x \times y = 1)) \\ &(\forall x)(\forall y)(\forall z)(x \times (y + z) = x \times y + x \times z) \quad 1 + 1 = 0. \end{aligned}$$

(ii): Posets with no maximal element

The language has $\Omega = \emptyset$ and $\Pi = \{\leq\}$ with arity 2. The theory has the following axioms:

$$\begin{aligned} &(\forall x)(x \leq x) \\ &(\forall x)(\forall y)((x \leq y \wedge y \leq x) \rightarrow x = y) \\ &(\forall x)(\forall y)(\forall z)((x \leq y \wedge y \leq z) \rightarrow x \leq z) \\ &(\forall x)(\exists y)(x \leq y \wedge x \neq y) \end{aligned}$$

Be alert to the difference between **maximal** elements and **maximum** elements.

(iii): Bipartite graphs

There are two correct answers.

(i) With a colour predicate:

The language has $\Omega = \emptyset$ and $\Pi = \{\sim, B\}$ of arities 2, 1. The theory has the following axioms:

$$\begin{aligned} &(\forall x)(\forall y)(x \sim y \longleftrightarrow y \sim x) \\ &(\forall x)(\forall y)(x \sim y \rightarrow (B(x) \wedge \neg B(y)) \vee (B(y) \wedge \neg B(x))) \end{aligned}$$

(ii) But you can also do it without the colour predicate, by asserting that there are no cycles of odd length. This needs infinitely many axioms. You might like to prove that bipartite graphs cannot be finitely axiomatised in the language of graph theory: it's a useful compactness exercise of the kind that you might meet in an exam

(iii) Actually there is a third correct answer which I hadn't considered, but which one of my students came up with. You could have a two-sorted language rather in the way that we might naturally have a two-sorted language for vector spaces. You have one set of variables for ranging over vertices, and another style of variable that ranges over colours. This is a much richer language and you can easily describe much more than just bipartite graphs. If you want a bipartite graph you have an axiom that says there are precisely two colours...

This method is of course extravagant, but the comparison between it and the method with a single colour predicate comes in useful later, with real vector spaces (part vii of this question). In

part vii the analogue of method three doesn't work: you have to do it by method one. But that's for later.

(iv) Indeed, I have now (may 2018) been shown a fourth answer, incorrect but very fertile (thank you ap888!!!). Evidently a graph is bipartite if you can adjoin two new vertices r and b , and edges to join each old vertex to precisely one of r and b in such a way that no two vertices connected to r (resp. b) are joined to each other. This *characterises* bipartite graphs but it does not *axiomatise* them... which is why it is not an answer to the question. Let us call a graph with two such vertices a wombat (you've got to call it *something*). Evidently wombats can be axiomatised in the language of graph theory. And evidently a graph is bipartite iff it is a subgraph of a wombat. This is rather like the fact that a ring is an integral domain iff it is a substructure of a field.

(iv): Algebraically Closed Fields

The language has $\Omega = \{+, \cdot, -, 0, 1\}$ with arities 2, 2, 1, 0, 0 and $\Pi = \emptyset$. The theory has the following axioms:

$$\begin{aligned} &(\forall x)(\forall y)(x + y = y + x) \\ &(\forall x)(\forall y)(\forall z)((x + y) + z = x + (y + z)) \\ &(\forall x)(x + 0 = 0) \\ &(\forall x)(x + (-x) = 0) \\ &(\forall x)(\forall y)(x \cdot y = y \cdot x) \\ &(\forall x)(\forall y)(\forall z)((x \cdot y) \cdot z = x \cdot (y \cdot z)) \\ &(\forall x)(x \cdot 1 = x) \\ &(\forall x)(x \neq 0 \rightarrow (\exists y)(x \cdot y = 1)) \\ &(\forall x)(\forall y)(\forall z)(x \cdot (y + z) = x \cdot y + x \cdot z) \\ &(\forall a_0) \dots (\forall a_n)(\exists x)(a_{n+1} \cdot x^{n+1} + a_n \cdot x^n + \dots + a_0 = 0). \end{aligned}$$

where the last axiom is understood as an axiom scheme ranging over all positive integers n .

(v): Groups of Order 60

The language has $\Omega = \{\cdot, ^{-1}, 1, g_1, g_2, \dots, g_{60}\}$ with arities 2, 1, 0, 0, \dots 0 and $\Pi = \emptyset$. The theory can be axiomatised as follows:

$$\begin{aligned} &(\forall x)(\forall y)(\forall z)(x \cdot (y \cdot z) = (x \cdot y) \cdot z) \\ &(\forall x)(x \cdot 1 = 1 \cdot x = x) \\ &(\forall x)(x \cdot x^{-1} = x^{-1} \cdot x = 1) \\ &(\forall x)(x = g_1 \vee x = g_2 \vee \dots \vee x = g_{60}) \\ &g_i \neq g_j \text{ for all } i \neq j \text{ (a scheme)} \end{aligned}$$

(vi): Simple Groups of Order 60

You might think you can use group presentations to axiomatise the theory of simple groups of order 60, but it's less than completely straightforward.

It's true that writing

$$\langle a^2 = b^3 = (ab)^5 = 1 \rangle$$

in some sense captures A_5 but it isn't enough by itself, since it appeals to the implicit information that no other equations hold, and that isn't first-order. Somehow you have to ensure that everything is in the group generated by a and b and you also have to ensure that no extra equations hold. The second point can be addressed by ensuring that there are 60 elements but that isn't much use unless we ensure that all those extra elements are denoted by words in a and b .

It may be that saying there are precisely 60 elements and every element is of order 2,3 or 5 and there are elements of all those orders is enough. I don't know enough group theory.

However something has emerged recently which is that, in every finite simple group, every element is a commutator. My guess is that the converse is true too, namely that every group where every element is a commutator is simple. If that's true then you add to the axioms of Group theory something to say that there are exactly 60 elements and

$$(\forall x)(\exists yz)(x = yzy^{-1}z^{-1})$$

Anyway the moral is that when you are trying to find a first-order axiomatisation of something that is obviously second-order you can—sometimes—cheat.

(vii): Real vector spaces

The language has $\Omega = \{+, -, 0\} \cup \{m_r : r \in \mathbb{R}\}$ with arities $2, 1, 0, 1, 1, \dots$ and $\Pi = \emptyset$. The theory has the following axioms:

$$\begin{aligned} &(\forall x)(\forall y)(\forall z)(x + (y + z) = (x + y) + z) \\ &(\forall x)(\forall y)(x + y = y + x) \\ &(\forall x)(x + 0 = x) \\ &(\forall x)(x + (-x) = 0) \\ &(\forall x)(\forall y)(m_r(x + y) = m_r(x) + m_r(y)) \\ &(\forall x)(m_{r+s}(x) = m_r(x) + m_s(x)) \\ &(\forall x)(m_{rs}(x) = m_r(m_s(x))) \end{aligned}$$

where the last three axioms are understood as axiom schemata ranging over all $r, s \in \mathbb{R}$.

Thank-you QY. All good, all true. However one might want to make the additional point that trying to axiomatise real vector spaces as a two-sorted theory doesn't work. It might seem natural to have two styles of variables with Latin letters for variables over vectors and Greek letter for variables over scalars. Indeed this is standard practice. One then adds the obvious axioms. The trouble with this is that, since the language is countable, the resulting theory will have countable models. What are the scalars in this countable model? They can't be the reals, co's the reals are uncountable. The scalars will form a field that is elementarily equivalent to the reals, but is countable. You *have* to do it the way QY does.

Of course this argument doesn't work for vector spaces over a *finite* field.

Question 9

I'm assuming that the reader has discovered the back-and-forth construction. I can't be bothered to explain it here, co's it's best done interactively in real time.

It is fairly easy to use the denseness of the rationals to show that every countable linear order can be embedded (in an order-preserving way) into \mathbb{Q} . Think of your countable total order as

the members of \mathbb{N} written in a funny order, and then find homes for the natural numbers one by one. That's OK but sadly it isn't quite enough, coz it goes only one way. You might next think "Suppose I have two countable dense linear orders ... I can embed each in the other—so I can then use Cantor-Bernstein!" That doesn't work, beco's Cantor-Bernstein works for *cardinals* not for linear order types—they're far too delicate. (After all, each of the two half-open intervals $(0, 1]$ and $[0, 1)$ embeds in the other but the two are not isomorphic.) So rather than build two embeddings separately, you *interleave* the two constructions in such a way that you construct a single isomorphism—a bijection.

Mind you, there actually *is* a version of Cantor-Bernstein for total orders, even tho' it is no use to us here. If A is iso to a terminal segment of B and B is iso to an initial segment of A then A and B are iso. ... Actually this is really a theorem about circular orders.

A follow-up thought...

Look at this once you've done sheet 4. Now that you have done ordinals and know what \aleph_1 is—the size of the set of countable ordinals—you might like to think about a generalisation of the fact that by a back-and-forth argument you can show that any two countable dense linear orders without endpoints are isomorphic. There is a theorem that says that any two dense linear orders of size \aleph_1 without endpoints are isomorphic (by a back-and-forth argument) as long as as they both satisfy a special extra condition. What is that extra condition?

Question 10

Easy to show that the theory of fields of characteristic 0 is axiomatisable. Merely add the scheme

$$\underbrace{1 + 1 + \dots + 1}_{p \text{ times}} \neq 0 \quad \text{for each } p$$

to the field axioms.

Slightly harder to show that it is not *finitely* axiomatisable. We exploit the following trivial fact:⁴ Suppose T is a theory with an infinite axiomatisation A such that no finite subset of A axiomatises T . Then T has no finite axiomatisation. For suppose it did. Let ϕ be the conjunction of the finite set of axioms. We have $A \vdash \phi$. Then, by compactness, we have $A' \vdash \phi$ for some finite $A' \subseteq A$. But this, by hypothesis, we do not have. Observe that the above axiomatisation of the theory of fields of characteristic 0 is an infinite axiomatisation no finite subset of which suffices so we can exploit the trivial fact.

There is a temptation to think that if the theory of fields of characteristic 0 has a finite axiomatisation then it has one in which the field axioms are separately itemised, so that the remaining axioms can be conjoined into a single axiom which in effect says "the field is of characteristic 0". Then you replace this axiom by its negation to obtain an axiomatisation of the theory of fields of positive characteristic, which of course is impossible. This can in fact be made to work, but it is not as straightforward as the proof i have just given. How can we be sure we can corral off the field axioms in this way? There is some work to do. Let our finite axiomatisation be the single formula ϕ . ϕ certainly implies the conjunction— F , say—of the field axioms. Now replace the single axiom ϕ with the two axioms $F \rightarrow \phi$ and F .

⁴I know it is trivial beco's i worked this out for myself when i was a mere philosophy student... a much lower lifeform than you, Dear Reader!

Are we now home and hosed? The candidate theory of fields of positive characteristic we obtain will be the field axioms F plus the negation of the remaining axiom $F \rightarrow \phi$. This negation is $F \wedge \neg\phi$, so this amounts to adding $\neg\phi$ as an axiom. Clearly no model \mathfrak{M} of $F \wedge \neg\phi$ can be a model of ϕ so \mathfrak{M} must be a field [beco's $\mathfrak{M} \models F$] and a field of positive characteristic. Converse? Let \mathfrak{M} be a field of positive characteristic. It's a model of F , because it's a field, but it can't be a model of ϕ beco's it isn't of characteristic 0. So $\{F, \neg\phi\}$ would be an axiomatisation of the theory of fields of positive characteristic [which we know to be impossible] so there really is no such ϕ .

Question 11

QY sez: “No Adjoin to the language a constant c and adjoin to the axioms of Peano arithmetic the sentences $0 < c$, $s(0) < c$, $s(s(0)) < c$, . . . to obtain a new theory S . Each finite subset of S has a model, so by compactness S has a model, which is of course infinite. By downward Löwenheim-Skolem, it has a countable model \mathfrak{M} . In \mathfrak{M} there is an element c which is greater than 0, $S(0)$, $S(S(0))$. . . but there is no such element in the standard model \mathbb{N} , so \mathfrak{M} is a nonstandard countable model of Peano arithmetic.”

Thanks for this QY, but classroom experience teaches me not leave it at that. Very well, so we have a model of arithmetic with an extra element. But it doesn't stop there. PA proves a whole lot of theorems saying that \mathbb{N} is closed under a lot of operations: $x \mapsto x^2$, $x \mapsto \lceil 22x/7 \rceil$, $x \mapsto \lceil \sqrt{x} \rceil$ and so on. It is probably quite helpful to think of our model as something containing 0 and c and *generated by them*. At its most basic it is a theorem of the arithmetic of \mathbb{N} , after all, that every number has a successor—and that every nonzero number has a predecessor—so we must have $c+1$ and $c-1$. This leads us to the conclusion that c belongs to a copy of \mathbb{Z} stuck on the end of \mathbb{N} . Only one copy. . . ? What about $\lceil 22c/7 \rceil$, $\lceil 355c/133 \rceil$. . . ? In fact a copy of \mathbb{Z} for every rational!

This has the striking (but as far as i know, useless) consequence that all countable nonstandard models of PA are isomorphic as ordered sets. So every countable nonstandard model of PA has order type $\mathbb{N} + \mathbb{Q} \cdot \mathbb{Z}$. You might think that you get *more* than \mathbb{Q} copies of \mathbb{Z} beco's of $\lceil \sqrt{c} \rceil$ but—as noted above, \mathbb{Q} is a maximal countable linear order type so you don't get any further copies of \mathbb{Z} by considering $\lceil \sqrt{c} \rceil$. Of course they aren't all isomorphic as structures for $+$ and \times —beco's arithmetic is incomplete.

I have just learnt the curious fact that every countable nonstandard model of PA is isomorphic to a proper initial segment of itself!

One point one sometimes has to make in this connection is that these wild and woolly things—the nonstandard naturals—living in the desolate marches beyond the standard naturals are absolutely **not** the same wild and woolly things living in the desolate marches beyond ω , namely the countable ordinals. This mistaken identification is a common consequence of over-enthusiastic fault-tolerant pattern matching by beginners.

Question 12

A group with an element of infinite order.

The language has $\Omega = (\cdot, {}^{-1}, 1, g)$ with arities 2, 1, 0, 0 and $\Pi = \emptyset$. The theory can be described by the following axioms:

$$\begin{aligned}
& (\forall x)(\forall y)(\forall z)(x \cdot (y \cdot z) = (x \cdot y) \cdot z) \\
& (\forall x)(x \cdot 1 = 1 \cdot x = x) \\
& (\forall x)(x \cdot x^{-1} = x^{-1} \cdot x = 1) \\
& \underbrace{g \cdot g \cdot \dots \cdot g}_{n \text{ times}} \neq 1 \text{ (for each } n \in \mathbb{N})
\end{aligned}$$

Can this be done purely in the language of groups? The answer Prof. Leader wants is ‘no’ and he is obviously correct, as we will see. However, Prof. Leader is a mere mortal (tho’ he might not appear to be, on cursory inspection) and the question contains a mistake. Since it is possible to axiomatise group theory just in the language with a single binary function symbol, you can go ahead and do it that way and—since you now no longer need the symbol ‘ e ’ to denote the unit—you can recycle that symbol to denote the element of infinite order! But that’s cheating, and the student who did it has been granted name suppression.

There now follows a proof of the impossibility of doing this in the language of groups, reconstructed from a recent conversation I had with Prof Leader.

The key is to find two groups one of which has an element of infinite order and the other does not, and yet the two groups are elementarily equivalent (indistinguishable by first-order expressions). To this end consider a group with elements of arbitrarily large finite order but no elements of infinite order. The group $\text{FSymm}(\mathbb{N})$ of permutations of \mathbb{N} that move only finitely many things will do nicely. Now consider the theory $T = \text{Th}(\text{FSymm}(\mathbb{N}))$ consisting of all the expressions in the language of group theory that hold in this group. This theory might not have a decidable set of axioms, but it doesn’t matter. What *does* matter—indeed is absolutely crucial—is that it is a **complete** theory. We now add a constant g to the language, and the obvious axioms $g^n \neq e$, for all $n \in \mathbb{N}$. Call the resulting theory T' . T' is clearly consistent by compactness and must have a model, which will be a group, call it G . G is a model of the complete theory $\text{Th}(\text{FSymm}(\mathbb{N}))$ and is therefore elementarily equivalent to $\text{FSymm}(\mathbb{N})$. But G has an element of infinite order and $\text{FSymm}(\mathbb{N})$ does not.

It doesn’t much matter that we took our group to be $\text{FSymm}(\mathbb{N})$. Any group with elements of arbitrarily large finite order but none of infinite order will do.

This works, and it’s very pretty, but it’s a bit *ad hoc*. Nathan Bowler points out to me that the additive group of the rationals mod 1 (“the rational circle”) has no element of infinite order (p/q is of order q) but the reals mod 1 (“the real circle”) has elements of infinite order. My guess is that these two groups are elementarily equivalent, and indeed that the inclusion embedding is elementary. By this we mean that, for any expression $\phi(\vec{x})$ in the language of groups, if $\phi(\vec{p})$ holds of some tuple \vec{p} in the additive group of the rationals mod 1, then it holds of the same tuple of rationals in the bigger group of reals mod 1. I might get round to writing out a proof. If it works (and I’m not making any promises) it would be a nicer proof than Prof Leader’s (although it’s much more involved) beco’s it is an introduction to a new technique.

Question 13

I tried to persuade Prof Leader that this question should be starred. He agrees that it’s hard, but he says it’s not *quite* hard enuff for a star.

The theory T has the following axioms:

$$\begin{aligned}
& (\forall x)(\forall y)(f(x) = f(y) \rightarrow x = y) \\
& (\forall y)(\exists x)(f(x) = y) \\
& \quad \quad \quad \underbrace{\hspace{1.5cm}}_{n \text{ times}} \\
& (\forall x)(\underbrace{f(f(f \cdots (x) \cdots))}_{n \text{ times}}) \neq x \text{ (for each } n \in \mathbb{N})
\end{aligned}$$

Any countable model \mathfrak{M} of T is a disjoint union of at most countably many f -cycles, all of which are of the form $\{\dots f^{-2}(x), f^{-1}(x), x, f(x), f^2(x), \dots\}$ for some x .

Imagine you are living in a world where there is nothing going on other than lots of points joined together by f edges, and all you can ever do is move along f edges (in either direction) from one point to another. What do you discover? By the end of time you have discovered that you are living on a copy⁵ of \mathbb{Z} . And that's *all* you have discovered: if the model contains another copy of the \mathbb{Z} -gon that you could have been on you never learn this fact. There is no way, in the given language, of saying that two vertices lie on distinct \mathbb{Z} -gons.

This is an informal picture and is definitely not a proof, but it might lead us to one.

I *think* that the model consisting of a single copy of \mathbb{Z} is what they call a **prime model**: it injects elementarily into all models of T . Presumably we use quantifier-elimination.

This could serve as an introduction to *Ehrenfeucht Games* but i can't go into that sort of detail here.

But there is a proof using only techniques available to you. (There must be, since this question isn't starred.) You observe that, altho' T can have nonisomorphic *countable* models (one, two or many copies of \mathbb{Z}), all its models of size 2^{\aleph_0} are isomorphic. This may not be immediately obvious. If \mathfrak{M}_1 and \mathfrak{M}_2 are two models both of size 2^{\aleph_0} then they both consist of 2^{\aleph_0} \mathbb{Z} -gons. (A detailed proof of this fact needs a little bit of AC but i'll spare you the details). So there is a bijection between the (set of) \mathbb{Z} -gons-in- \mathfrak{M}_1 and the (set-of) \mathbb{Z} -gons-in- \mathfrak{M}_2 . This isn't *quite* a bijection between M_1 and M_2 , but we are nearly there. All we have to do is pick, for each pair of a- \mathbb{Z} -gon-in- \mathfrak{M}_1 -with-a- \mathbb{Z} -gon-in- \mathfrak{M}_2 , a digraph isomorphism between the two \mathbb{Z} -gons, and take the union of all those isomorphisms. This union will be an isomorphism between \mathfrak{M}_1 and \mathfrak{M}_2 . If T were not complete we would be able to find ϕ such that $T \cup \{\phi\}$ and $T \cup \{\neg\phi\}$ were both consistent. Add 2^{\aleph_0} constants and deduce (by compactness) that $T \cup \{\phi\}$ and $T \cup \{\neg\phi\}$ both have models of size at least 2^{\aleph_0} . Indeed (by downward Skolem-Löwenheim) they must both have models of size *precisely* 2^{\aleph_0} . These models would have to be nonisomorphic beco's one of them believes ϕ and the other believes $\neg\phi$. But they are both models of T so they are isomorphic.

Instead of 2^{\aleph_0} one can use \aleph_1 . Students would be unlikely to try doing it that way beco's \aleph_1 is a mysterious phobic object for them. But it works better. In particular one does not need AC, at least not after the use of AC to prove upward Skolemheim to show that there is a model of size \aleph_1 . What does a model of T of size \aleph_1 look like? Lots of copies of \mathbb{Z} of course, but precisely how many? The set of copies of \mathbb{Z} is a surjective image of a set of size \aleph_1 and so is of cardinality $\leq \aleph_1$. The copies of \mathbb{Z} have a global wellordering, so the size of their union (which is \aleph_1) is \aleph_0 times something; it can only be \aleph_1 .

Sometimes students can be *soooo* annoying. The point of this question (as you have probably guessed by now) is to direct your attention to theories that are categorical in some *uncountable* cardinal. However there is a way of answering this question that doesn't exploit this possibility,

⁵Actually it's not really \mathbb{Z} beco's \mathbb{Z} has additive and multiplicative structure, which this thing hasn't. It's really just a digraph. One might call it the **\mathbb{Z} -gon**.

and some of you found it. That was not in the script at all. Grrr! Suppose $T \not\models \phi$ and $T \not\models \neg\phi$. Add countably many constants to the language of T , and add axioms to $T \cup \{\phi\}$ and to $T \cup \{\neg\phi\}$ to say that the denotations of these constants all belong to different \mathbb{Z} -gons. These two theories (call them T_1 and T_2) both have countable models by downward Skolemheim. What can countable models of these theories look like? They must consist of precisely \aleph_0 \mathbb{Z} -gons infinitely many of which have a distinguished element in each \mathbb{Z} -gon. There's no way of compelling every \mathbb{Z} -gon to have a distinguished element, so this doesn't completely wrap things up for us. However, you weren't supposed to do it that way anyway!

Question 14

I am going to keep to my practice of humouring Prof. Leader's desire to prevent answers to starred questions leaking out. For my part i know what the answer is supposed to be, but I have to confess that i have never worked through a proof. I suspect it isn't hard.

Sheet 4

Question 1

Challenge: deduce the axiom of empty set from the axiom of infinity.

The morally correct way to do this is to observe that the axiom of infinity has the form “there is a set with a special property”. If there is even one set, then—as long as we have separation—there will be an empty set, since the subsets consisting of all those elements of the set that are not equal to themselves will be a set by separation.

Now the axiom of infinity can also come in the form “There is a successor set”, or

$$(\exists x)(\emptyset \in x \wedge (\forall y)(y \in x \rightarrow y \cup \{y\} \in x))$$

In the presence of the axiom scheme of replacement this can be deduced from the bare assertion there is an infinite set (a set not the same size as any proper subset of itself) but the axiom is often taken in this more specialised form because it makes it easy to give immediately an implementation of arithmetic. Fair enough. However, this muddies the waters slightly, in that it enables us to give a different proof of the existence of the empty set. People sometimes say that the axiom of infinity *presupposes* the existence of the empty set, but that’s not quite right. Let’s get this 100% straight. The axiom in the form “there is a successor set” says

$$(\exists x)((\exists e \in x)(\forall w)(w \notin e) \wedge (\forall y)(y \in x \rightarrow y \cup \{y\} \in x)) \quad (2.2)$$

I have written out the ‘ $\emptyset \in x$ ’ bit in primitive notation so we can be sure that there are no tricks being played.

The expression (2.2) is of the form

$$(\exists x)(p \wedge F(x))$$

where p is $(\exists e \in x)(\forall w)(w \notin e)$ and $F(x)$ is $(\forall y)(y \in x \rightarrow y \cup \{y\} \in x)$. Anything of the form $(\exists x)(p \wedge F(x))$ is going to imply $(\exists x)p$, namely

$$(\exists x)(\exists e \in x)(\forall w)(w \notin e)$$

whence

$$(\exists e)(\forall w)(w \notin e)$$

which says that there is an empty set, which is what we wanted.

Last part, deducing the axiom (scheme) of separation from the axiom (scheme) of replacement.

If replacement allows you to use *partial* functions it’s easy. If you are only allowed *total* functions then you want Phil Connell’s trick (tidied up by me to make it constructive):

Define $f(x)$ to be $\{y : y = x \wedge \phi(y)\}$. This has the effect that f sends to their singletons the things you want to keep, and sends everything else to the empty set. Then $\bigcup f$ “ W is $\{x \in W : \phi(x)\}$ ”. Observe that this is constructive.

The other way (preferable in certain circumstances) is to say: *either* there is nothing in W which has ϕ (in which case the set we want is the empty set, and we have an axiom for that) or

there is an $x \in W$ s.t. $\phi(x)$. For any such x we can define a function f which sends $y \in W$ to y as long as $\phi(y)$, and sends y to x o/w.

What is there to choose between these two proofs? Phil Connell's proof uses the axiom of sumset, but the second method uses excluded middle. (It uses it *twice*; once in the case split, and again in the second of the two cases, testing whether or not $\phi(y)$).

Question 2

Two applications of power set to \emptyset gives you $\{\{\emptyset\}, \emptyset\}$ which we then whack with the function class

$$(u = \{\emptyset\} \wedge v = x) \vee (u = \emptyset \wedge v = y)$$

which will give us the pair $\{x, y\}$.

Question 3

"Write down sentences in the language of set theory to express the assertions that, for any two sets x and y , the product $x \times y$ and the set y^x of all functions from x to y exist. You may assume that your pairs are Wiener-Kuratowski."

If you use Wiener-Kuratowski pairs then $\langle x, y \rangle = \{\{x\}, \{x, y\}\}$ and is a subset of $\mathcal{P}^2(\{x, y\})$. Similarly $x \times y$ is a subset of the power set a couple of times of $x \cup y = \bigcup \{x, y\}$. Clearly the set of functions from x to y can be obtained in the same way.

What if you want to establish that these things are sets without knowing what your pairing function is? Imagine the following situation: i want $X \times Y$ and i know that there is a set-theoretic construct $\langle x, y \rangle$, tho' i don't know what it is and i'm not allowed to assume anything other than that it is there and is available. We do the following: fix $y \in Y$ and consider the function class that sends x to $\langle x, y \rangle$. The image of X in this function exists by replacement and it is of course $X \times \{y\}$. So $X \times \{y\}$ exists for all y . Now consider the function class that sends y to $X \times \{y\}$. The image of Y in this function exists by replacement and is $X \times Y$.

So: if we have replacement we can prove that $X \times Y$ exists *whatever implementation of pairing-with-unpairing we use*. You might like to prove the converse: if $X \times Y$ always exists for all implementations of pairing-with-unpairing then replacement follows.

Question 4

Neither direction works.

$\in\{\{\emptyset\}\}$ is the empty relation, and therefore transitive, but $\{\{\emptyset\}\}$ is not a transitive set.

For the other direction, consider

$\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$ is a transitive set, but $\in\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$ is not a transitive relation.

I think the danger in this second part is to reason "Suppose $z \in y \in x$; then, since \in is transitive, we infer $z \in x$ ". This doesn't work. It assumes that $\in \cup \{x\}$ is transitive which sounds the same but isn't.

Question 5

Use \in -induction to show that the only \in -automorphism definable by a function-class is the identity.

Assume that f is an \in -automorphism, which is to say $f(x) = f''x$ for all x . But “every member of x is fixed by f ” is just $x = f''x$, whence $f(x) = x$. So x is fixed as long as all its members are; \in -induction does the rest.

Discussion

Notice we could be doing one of two inductions here, and it just might matter which. We could specialise to a given \in -automorphism f and then prove by \in -induction that everything is fixed by it; alternatively we could prove by \in -induction that “ x is fixed by all \in -automorphisms”. This second formulation has a universal quantifier ranging over *classes* and that raises technical difficulties for ZF since it involves quantifying over proper classes. This is something that you don’t want to think about, at least not yet.

Question 6

(i) Determine the rank of the set \mathbb{R} of real numbers. [You may assume that a real number is an ordered pair of subsets of \mathbb{Q} (a Dedekind section), that a rational number is an equivalence class of ordered pairs of integers, and so on.]

(ii) Show that there is a subset of \mathbb{R} which (under the restriction of the usual ordering on \mathbb{R}) is order-isomorphic to $\omega + \omega$.

(iii) Show that all the axioms of ZF except for the scheme of Replacement hold in $V_{\omega+\omega}$. Why can we deduce from (ii) that Replacement does **not** hold?

Discussion

This question makes several points. One of them is the point that there are lots of ways of implementing \mathbb{Z} , \mathbb{Q} , \mathbb{R} etc as sets; another is that—mathematically at least—it doesn’t much matter which one you use. The other is to get you to do some set-theoretic calculations—computing the ranks of particular sets.

One should start with a warning, the (set-theoretic) rank of a set equipped with an ordering cannot be computed from the order-type of the ordering: it’s a property of the set, not of any ordering of it. And again, it’s nothing to do with cardinality either, or very little. There are small sets (*singletons* indeed) of arbitrarily high rank.

So the rank of a mathematical object implemented as a set is not a mathematical invariant of that object. This comes as a surprise to many students, so it is worth making a big song-and-dance about it. Admittedly it is true that the *minimum possible* rank of an implementation of that object is a mathematical invariant of that object [I think Prof Leader mentioned it under the monnicker “essential rank”] but it’s a curiously uninteresting one, being controlled entirely by cardinality. You cannot implement \mathbb{R} as an object of rank ω beco’s there are too few things of lower rank for all the reals to be implemented by those things of lower rank. There are uncountably many reals but only countably many things of finite rank. This cardinality consideration is the only constraint.

Actually in parts of the set theoretic literature reals are taken to be functions from \mathbb{N} to \mathbb{N} , things sometimes called *set theorists' reals*. They have rank ω —which is best possible, for the reason given above. (They're something to do with continued fractions.)

So: pick an implementation, and compute the ranks of the sets you end up with. For bonus points, pick more than one implementation, and compute all of them! If you know what p -adic numbers are, compute their rank too. (The p -adics are the completion of \mathbb{Q} w.r.t the p -adic metric. How are you to think of the completion set-theoretically?)

For (iii) observe that the fact that you can find a subset of \mathbb{R} that can be wellordered to length $\omega + \omega$ has the following ramifications. If replacement held in $V_{\omega+\omega}$ then $V_{\omega+\omega}$ would contain, for every wellordering in $V_{\omega+\omega}$ the corresponding von Neumann ordinal. It is easy to check that the rank of the von Neumann ordinal α is α itself, which means that $V_{\omega+\omega}$ cannot contain any ordinal from $\omega + \omega$ onwards. So replacement fails. In general V_α will contain at least some wellorderings that are far too long for their von Neumann ordinals to be in V_α . It happens only rarely that α “catches up” with the ordinals in V_α .

In contrast, the collection H_κ of sets hereditarily of size less than κ is practically guaranteed to be a model of replacement, as follows. Suppose $X \in H_\kappa$, and $f : H_\kappa \rightarrow H_\kappa$. Then $f''X$ is a subset of H_κ . How big is it? It's a surjective image of thing of size $< \kappa$. We want it to be of size $< \kappa$ itself. So all we need is a surjective image of something of size less than κ is itself of size $< \kappa$. This is certainly true if κ is an aleph, and even in many cases when it isn't. So certainly if κ is an aleph then H_κ is a model of replacement.

Question 7

The key to this question is induction, both structural and wellfounded.

One direction is easy: you prove by induction on n that everything in V_n is in HF . For the other direction you have to use \in -induction to show $HF \subseteq V_\omega$. The property $\phi(x)$ you prove by \in -induction is “ $x \in HF \rightarrow x \in V_\omega$ ”.

If foundation fails then potentially a Quine atom⁶ is a counterexample to the inclusion.

Question 8

You can explain why $V_{\omega+\omega} \not\models$ replacement without using the axiom of sumset. Consider the function class $n \mapsto V_{\omega+n}$. Replacement would make the image of \mathbb{N} in this function class—namely $\{V_{\omega+n} : n \in \mathbb{N}\}$ —into a set of the model, and it can't be, beco's it is of rank $\omega + \omega$.

Codicil to questions 7 and 8

Take care when asking yourself whether or not an axiom is true in a structure. Yer typical set theoretic axiom states that the universe is closed under some operation (as it might be power set, or sumset). **Saying that a structure is a model for that axiom is not the same as saying that it's closed under the corresponding operation.**

See the discussion of question 12 of this sheet.

⁶A Quine atom is a set $x = \{x\}$.

Question 9

You all think you know that $|\mathbb{R}| = 2^{\aleph_0}$ and you're right of course but finding a bijection between \mathbb{R} and $\mathcal{P}(\mathbb{N})$ is not an *absolute* doddle. It can do you no harm to track one down.

Jonathan Holmes has the cutest proof of this fact known to me (I have doctored this from his answer to this sheet). Every real number has a unique representation as an ω -string of 0's and 1's *containing arbitrarily late 0s*. The italicised condition removes duplicate representations of dyadic rationals. Each such string corresponds to a set (of addresses where the string has 1s) whose complement is infinite. How many subsets of \mathbb{N} are there whose complement in \mathbb{N} is infinite? Well, there are \aleph_0 subsets of \mathbb{N} that do not satisfy this condition, so we are looking at $\mathcal{P}(\mathbb{N})$ minus a countable set. You then use Bernstein's lemma to show that any such set has cardinality precisely 2^{\aleph_0} .

To return to Q9, the clever way to prove this is to observe that the function $F : (\mathbb{R} \rightarrow \mathbb{R}) \rightarrow (\mathbb{R} \rightarrow \mathbb{Q})$ that takes a continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$ and returns $f \upharpoonright \mathbb{Q}$ (its restriction to \mathbb{Q}) is injective. Thus F injects the set of continuous functions $\mathbb{R} \rightarrow \mathbb{R}$ into the set of functions $\mathbb{Q} \rightarrow \mathbb{R}$. This latter set is of size $(2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0}$.

This illustrates a general phenomenon. If the set you are trying to compute the size of is naturally a subset of the n -times power set of \mathbb{N} or \mathbb{R} etc then its size will be an iterated exponent of \aleph_0 . As I remarked on p. 92, sizes of quotients can be very complicated. How many wellorders are there of \mathbb{N} , *up to isomorphism*? The answer is \aleph_1 , which may or may not be equal to 2^{\aleph_0} . Contrast the two questions:

How many total orders of \mathbb{N} are there whose automorphism group is transitive on singletons? (A)

How many countable order types are there whose automorphism group is transitive on singletons? (B)

The answer to (A) is obviously 2^{\aleph_0} . The set in (B) is a quotient of the set in (A) and calculating its size is a hard task.

Question 10

Consider the sequence $S = \omega, \omega_\omega, \omega_{\omega_\omega} \dots$ of von Neumann ordinals. It's supremum (union) is obviously going to be a fixed point. However, this question is on a *Set Theory* sheet not an *Ordinals* sheet, so you should be thinking quite hard about how we use the resources of set theory to prove that there really is a wellordering of this length. So we should be asking: how do we know the ordinals stretch that far? The proof is a long road...

For a start, how do we even know that the sequence is even there at all for us to take its sup? Clearly we are going to need an instance of the axiom scheme of replacement. Whack \mathbb{N} with the function class that sends n to $\omega_{\dots\omega}$ with n dots. How do we know that this function is defined for all natural numbers? Probably by induction on naturals. Start with ω_ω . How do we know that there is a wellordering of this length? Well, ω_ω is the sup of $\omega, \omega_1, \omega_2 \dots$, and we know that each of these exists by Hartogs' lemma. Then we obtain ω_ω by replacement again. (And it is known that you need replacement to prove the existence of wellorderings that long). And how are you going to get from ω_ω to ω_{ω_ω} ?

Question 11

The first sentence contains an ellipsis. He doesn't really mean you to show that there is no surjection from \aleph_n onto \aleph_{n+1} because you don't know what sets these cardinals are. (In fact in both the von Neumann implementation and the Scott's trick implementation this allegation is, as it happens, true). What he means is that you should show that there is no surjection from **a set of size** \aleph_n onto **a set of size** \aleph_{n+1} . This is pretty easy, even if you aren't allowed choice. It follows from the simple observation that if there is a surjection $X \twoheadrightarrow Y$, and X can be wellordered, then there is an injection $Y \hookrightarrow X$.

The rest of the question is the **Jordan-König theorem**, which says that, if you have two families $\{A_i : i \in I\}$ and $\{B_i : i \in I\}$ of sets, where, for each $i \in I$, there is no surjection of A_i to B_i then the union $\bigcup_{i \in I} A_i$ of the A s does not map onto the product $\prod_{i \in I} B_i$ of the B s. The idea is as follows. Suppose

$$f : \bigcup_{i \in I} A_i \rightarrow \prod_{i \in I} B_i$$

We will show that f is not surjective. We have to exploit somehow the circumstance that no function $A_i \rightarrow B_i$ can be surjective, so the obvious thing to do is extract from f a family of functions $f_i : A_i \rightarrow B_i$, and trade on the fact that f_i is not surjective. To this end we declare that $f_i(a)$ is to be $f(a)$ applied to i . By assumption $B_i \setminus f_i[A_i]$ is nonempty, so let $g \in \prod_{i \in I} B_i$ be a function defined so that $g(i) \in B_i \setminus f_i[A_i]$ at each i . I'll leave it to you to show that g is not in the range of f .

Actually you don't need the *whole* of the Jordan-König theorem, and I think what Dr Russell wants you to do is use that line of thinking to show that there is no surjection of $A \twoheadrightarrow (\mathbb{N} \rightarrow A)$ if $|A| = \aleph_\omega$... using the fact that \aleph_ω is the sup of the \aleph_n , with $n \in \mathbb{N}$.

There is a proof of Jordan-König in *Logic, Induction and Sets*. The full general version needs choice⁷ but if you merely want to show that $2^{\aleph_0} \neq \aleph_\omega$ then you don't. (It will do you no harm to think about why not)

Finally you might like to check your understanding of this situation by proving that 2^{\aleph_0} cannot be equal to \aleph_α if the cofinality of α is ω .

Question 12

This is all about absolute properties *versus* non-absolute properties. Ha Thu Nguyen gives a very simple illustration ... If $\mathfrak{M} = \langle M, \in \rangle$ is a model of ZF then it doesn't think that its carrier set M is a set, but we can see from outside that it is. Tim Talbot puts it very well: the countable model is a *Tardis*!

Question 13

A union of countably many countable sets cannot have size \aleph_2 .

It's easier if you attempt to prove that

⁷Is *equivalent* to choice indeed: think ... what happens if the product of nonempty sets is not reliably nonempty?

if every A_α is of cardinality \aleph_γ at most, **and** $\bigcup_{\alpha < \omega_\beta} A_\alpha$ can be wellordered
then $|\bigcup_{\alpha < \omega_\beta} A_\alpha| \leq \max\{\aleph_\beta, \aleph_\gamma\}$.

The point being that if $\bigcup_{\alpha < \omega_\beta} A_\alpha$ is wellordered then you can use the restrictions of that wellordering to the various A_i to exploit the fact that $\aleph_\gamma \cdot \aleph_\gamma = \aleph_\gamma$.

Question 14

Be careful how you read this question. You certainly can't add $\mathbb{1}$ to the ring of V if \cdot is \cap and $+$ is Δ , because the $\mathbb{1}$ would have to be V and then one application of separation will give you the Russell class as a set and you'd get Russell's paradox. The question is: are there other ring structures you can impose on V which give you a $\mathbb{1}$?

Actually there are several. Prof Leader's favoured solution is to consider, say, the ring $\mathbb{Z}[V]$ of polynomials over \mathbb{Z} with one variable for every set. A bit of Cantor-Bernstein makes this the same size as V so you can copy the ring structure over to V . That looks to me like an unmathematical trick, but there may be attractive features that I have missed. In fact, knowing Professor Leader, there probably are.

Is there another way? If you know about Conway numbers (and if you don't you should) you will recall that the ordinals can be given the structure of a field of characteristic 2. You might think you can copy this over to V . However the project of copying-the-field-structure needs a bijection between V and On and the axiom that says that there is such a bijection is the axiom of *global* choice, and that axiom is strictly stronger than the axiom you have been given (which is the axiom of *local* choice).

Both these solutions are nice, but there is one that is nicer still. Nicer because it provides a $\mathbb{1}$ while retaining \cdot as \cap and $+$ as Δ . To do that of course you need a different membership relation. This is the device of *Church-Oswald models*. See stuff about them on my home page: <http://www.dpmms.cam.ac.uk/~tf/church2001.pdf>). In brief you define a new membership relation as follows. There is a bijection (call it k) between V and $V \times \{0, 1\}$. We then define

$$x \in_{new} y \text{ iff } \begin{cases} \text{snd}(k(y)) = 0 \text{ and } x \in \text{fst}(k(y)) & \text{or} \\ \text{snd}(k(y)) = 1 \text{ and } x \notin \text{fst}(k(y)) \end{cases}$$

This makes the universe into a boolean ring—in fact a boolean algebra where \vee is \cup , \wedge is \cap and complementation is set complementation. (Check: $k^{-1}\langle \emptyset, 1 \rangle$ becomes the universal set)

CO models are my favoured solution, being natural and interesting structures—models for natural and interesting set theories, indeed.

Question 15

This was once an exam question for my Part III Computability and Logic course. It's not *terribly* hard (the question about evil total orders from sheet 2 is a lot harder) and I know of at least two ways of doing it (there may be more) but the details are quite fiddly.

In keeping with my project to keep Prof Leader sweet I am going to humour him and not supply an answer.

However, unlike the other starred questions this is one with great mathematical significance.

Chapter 3

tf's example sheets for 2016/7

Set Theory and Logic, Michaelmas 2016, Sheet 1: Ordinals and Induction

Questions marked with a '*' may be skipped by the nervous.

(i)

Write down subsets of \mathbb{R} of order types $\omega + \omega$, ω^2 and ω^3 in the inherited order.

One of my students came up with $\{1 - 1/n : n \in \mathbb{N}\} \cup \{10 - 1/n : n \in \mathbb{N}\}$. Why that rather than $\{1 - 1/n : n \in \mathbb{N}\} \cup \{2 - 1/n : n \in \mathbb{N}\}$, i wondered ...? His answer is the range of an order-preserving map from the ordinals below $\omega + \omega$ into \mathbb{R} . My preferred answer is the range of a *continuous* order-preserving map from the ordinals below $\omega + \omega$ into \mathbb{R} .

ω^2 is not that hard: $\{n - 1/m : n, m \in \mathbb{N}\}$, but ω^3 requires a bit of work. Fortunately most of you were up to it. The key observation is that, in each copy of ω , the gap between the m th and the $m + 1$ th point is $\frac{1}{m(m+1)}$ wide, so if you want to squeeze an extra copy of ω in there you do

$$\left\{n - \frac{1}{m} - \frac{1}{km(m+1)} : n, m, k \in \mathbb{N}\right\}$$

Test your comprehension by doing ω^4 in the same style.

(ii)

Which of the following are true?

- (a) $\alpha + \beta$ is a limit ordinal iff β is a limit ordinal;
- (b) $\alpha \cdot \beta$ is a limit ordinal iff α or β is a limit ordinal;
- (c) Every limit ordinal is of the form $\alpha \cdot \omega$;
- (d) Every limit ordinal is of the form $\omega \cdot \alpha$.

For these purposes 0 is a limit ordinal.

(iii)

Consider the two functions $On \rightarrow On$: $\alpha \mapsto 2^\alpha$ and $\alpha \mapsto \alpha^2$. Are they normal?

(iv)

Prove the converse to lemma ??: if $\langle X, <_X \rangle$ is a total order satisfying “every subordering is isomorphic to an initial segment” then it is a wellordering.

(v)

What is the smallest ordinal you can not embed in the reals in the style of question (i)?

(vi)

Prove that every [nonzero] countable limit ordinal has cofinality ω . What about ω_1 ?

(vii)*

Recall the recursive definition of ordinal exponentiation:

$$\alpha^0 = 1; \alpha^{\beta+1} = \alpha^\beta \cdot \alpha, \text{ and } \alpha^{\sup(B)} = \sup(\{\alpha^\beta : \beta \in B\}).$$

Ordinal addition corresponds to disjoint union [of wellorderings], ordinal multiplication corresponds to lexicographic product, and ordinal exponentiation corresponds to ...? Give a definition of a suitable operation on wellorderings and show that your definition conforms to the spec: $\alpha^{\beta+\gamma} = \alpha^\beta \cdot \alpha^\gamma$.

(viii)

Let $\{X_i : i \in I\}$ be a family of sets, and Y a set. For each $i \in I$ there is an injection $X_i \hookrightarrow Y$. Give an example to show that there need not be an injection $(\bigcup_{i \in I} X_i) \hookrightarrow Y$. But what if the X_i are nested? [That is, $(\forall i, j \in I)(X_i \subseteq X_j \vee X_j \subseteq X_i)$.]

(ix)

Prove that every ordinal of the form ω^α is **indecomposable**: $\gamma + \beta = \omega^\alpha \rightarrow \gamma = \omega^\alpha \vee \beta = \omega^\alpha$.

(x)

Show that an arbitrary intersection of transitive relations is transitive. The **transitive closure** R^* (sometimes written ‘ $tr(R)$ ’) is the \subseteq -least transitive relation $\supseteq R$.

Let $\langle X, R \rangle$ be a wellfounded binary structure, with rank function ρ . Prove that $(\forall x \in X)(\forall \alpha < \rho(x))(\exists y)(\rho(y) = \alpha)$.

Of course you do this by R -induction

[A later—perhaps preferable—version of this question...]

Let $\langle X, R \rangle$ be a wellfounded binary structure, with rank function ρ . Prove that $(\forall x \in X)(\forall \alpha < \rho(x))(\exists y \in X)(\rho(y) = \alpha)$.

(xi)

Let $\{X_i : i \in \mathbb{N}\}$ be a nested family of sets of ordinals.

- (a) Give an example to show that the order type of $\bigcup_{i \in \mathbb{N}} X_i$ need not be the sup of the order types of the X_i .
- (b) What condition do you need to put on the inclusion relation between the X_i to ensure that the order type of $\bigcup_{i \in \mathbb{N}} X_i$ is the sup of the order types of the X_i ?
- (c) Show that the ordered set of the rationals can be obtained as the union of a suitably chosen ω -chain of some of its finite subsets.

The point is that any structure whatever can be obtained as a direct limit (“colimit”) of its finitely generated substructures.

(xii)

Using the uniqueness of subtraction for ordinals, and the division algorithm for normal functions, show that every ordinal can be expressed uniquely as a sum

$$\omega^{\alpha_1} \cdot a_1 + \omega^{\alpha_2} \cdot a_2 + \cdots \omega^{\alpha_n} \cdot a_n$$

where all the a_i are finite, and where the α_i are strictly decreasing.

(xiii)

Let f be a function from countable [nonzero] limit ordinals to countable ordinals satisfying $f(\alpha) < \alpha$ for all (countable limit) α . (f is “pressing-down”.) Can f be injective?

Duplicated below.

Set Theory and Logic, Michaelmas 2016, Sheet 2: Posets

‘+’ signifies a question you shouldn’t have trouble with; ‘☹’ means what you think it means.

(i)

(a) For $n \in \mathbb{N}$, how many antisymmetrical binary relations are there on a set of cardinality n ? How many binary relations satisfying trichotomy: $(\forall xy)(R(x, y) \vee R(y, x) \vee x = y)$? How are your two answers related?

(b) How many symmetric relations and how many antisymmetric trichotomous relations are there on a set of cardinality n ? How are your two answers related?

(c) Contrast (a) and (b)

(ii)

Consider the set of equivalence relations on a fixed set, partially ordered by \subseteq . Show that it is a lattice. Must it be distributive? Is it complete?

It is complete but not distributive.

(iii)

Cardinals: Recall that $\alpha \cdot \beta$ is $|A \times B|$ where $|A| = \alpha$ and $|B| = \beta$. Show that a union of α disjoint sets each of size β has size $\alpha \cdot \beta$. Explain your use of AC.

(iv)

Let $\langle A, \leq \rangle$ and $\langle B, \leq \rangle$ be total orderings with $\langle A, \leq \rangle$ isomorphic to an initial segment of $\langle B, \leq \rangle$ and $\langle B, \leq \rangle$ isomorphic to a terminal segment of $\langle A, \leq \rangle$. Show that $\langle A, \leq \rangle$ and $\langle B, \leq \rangle$ are isomorphic.

This is a sleeper for Banach-Tarski. Look at Wagon's book.

(v)

(Mathematics Tripos Part II 2001:B2:11b, modified).

Let U be an arbitrary set and $\mathcal{P}(U)$ be the power set of U . For X a subset of $\mathcal{P}(U)$, the **dual** X^\vee of X is the set $\{y \subseteq U : (\forall x \in X)(y \cap x \neq \emptyset)\}$.

1. Is the function $X \mapsto X^\vee$ monotone? Comment.

It's obviously monotone decreasing wrt \subseteq . But we can say a bit more than that: it's continuous: $(\bigcup\{a_i : i \in I\})^\vee = \bigcap\{(a_i)^\vee : i \in I\}$. We'd better prove this!

2. By considering the poset of those subsets of $\mathcal{P}(U)$ that are subsets of their duals, or otherwise, show that there are sets $X \subseteq \mathcal{P}(U)$ with $X = X^\vee$.

It is chain-complete. Let $\{a_i : i \in I\}$ be a chain, with I an ordered set, so that $i \leq_I j \rightarrow a_i \subseteq a_j$. Consider the chain dual to this, namely $\{(a_i)^\vee : i \in I\}$. We want the dual to $\bigcup\{a_i : i \in I\}$ to be $\bigcap\{(a_i)^\vee : i \in I\}$. First we prove that $\bigcup\{a_i : i \in I\} \subseteq \bigcap\{(a_i)^\vee : i \in I\}$. To establish this we need to show that, for all $i, j \in I$, $a_i \subseteq (a_j)^\vee$. If $i <_I j$ then $a_i \subseteq a_j \subseteq (a_j)^\vee$; if $j <_I i$ then $a_j \subseteq a_i$ whence $(a_i)^\vee \subseteq (a_j)^\vee$, but we have $a_i \subseteq (a_i)^\vee$ giving $a_i \subseteq (a_j)^\vee$ as desired.

By part (1) we can now infer that the sup $(\bigcup\{a_i : i \in I\})^\vee$ of the chain is a subset of its dual $\bigcap\{(a_i)^\vee : i \in I\}$.

3. $X^{\vee\vee}$ is clearly a superset of X , in that it contains every superset of every member of X . What about the reverse inclusion? That is, do we have $Y \in X^{\vee\vee} \rightarrow (\exists Z \in X)(Z \subseteq Y)$?

4. Is $A^{\vee\vee\vee}$ always equal to A^\vee ?

If you think of \forall and \exists as sets, then \forall is the singleton of the domain, and \exists is the set of nonempty subsets of the domain, and these two are dual to each other in the sense of the question. This is of course the motivation.

For the converse, suppose that $Y \in A^{\vee\vee}$ but Y is not a superset of anything in A . Then consider $\bigcup\{x \setminus Y : x \in A\}$. It's clearly in A^\vee since it meets every $x \in A$ but it is disjoint from Y , so Y cannot be in $A^{\vee\vee}$. Thus $A^{\vee\vee}$ is the closure of A under superset.

For the rider (“Is $A^{\vee\vee\vee}$ always equal to A^\vee ?”) reflect that we have shown that $B \subseteq B^{\vee\vee}$ so we must have $A^\vee \subseteq A^{\vee\vee\vee}$. For the reverse inclusion suppose $Y \in A^{\vee\vee\vee}$. That is to say, Y meets everything in $A^{\vee\vee}$. How does this differ from meeting everything in A ? Not one whit, because the only difference between things in A and things in $A^{\vee\vee}$ is that any new things that appear in $A^{\vee\vee}$ are supersets of things in A , and if you meet every superset of y (y a member of A) then you meet y . This proves the inclusion.

This is (distantly) related to the fact that: although constructive logic does not obey double negation ($\neg\neg p$ is not reliably the same as p) it does obey *triple* negation: $\neg\neg\neg p = \neg p$.

(vi)

Use Zorn’s Lemma to prove that

- (i) every partial ordering on a set X can be extended to a total ordering of X ;
- (ii) for any two sets A and B , there exists either an injection $A \hookrightarrow B$ or an injection $B \hookrightarrow A$.

(vii)

(Tripos IIA 1998 p 10 q 7)

Let $\langle P, \leq_P \rangle$ be a chain-complete poset with a least element, and $f : P \rightarrow P$ an order-preserving map. Show that the set of fixed points of f has a least element and is chain-complete in the ordering it inherits from P . Deduce that if f_1, f_2, \dots, f_n are order-preserving maps $P \rightarrow P$ which commute with each other (i.e. $f_i \circ f_j = f_j \circ f_i$ for all i, j), then they have a common fixed point. Show by an example that two order-preserving maps $P \rightarrow P$ which do not commute with each other need not have a common fixed point.

Discussion

[PTJ says: The first part is similar to part (ii) of question 4. For the second, first show that each of f_2, \dots, f_n maps the set of fixed points of f_1 into itself, and then use induction on n . (It’s worth pointing out the similarity with the technique—which they should have seen before—for finding a common eigenvector of a family of commuting endomorphisms.) The simplest counterexample for the last part is to take two distinct constant maps].

If $f_1 \dots f_n$ are order-preserving maps $P \rightarrow P$ which commute with each other then their set of common fixed points is chain-complete and has a least element.

Proof: by induction on n . Base case easy! If $n > 1$ by induction hypothesis the set F of common fixed points of $f_1 \dots f_{n-1}$ is chain complete and has a least element. Since f_n commutes with all f_k with $k < n$, f_n sends F into itself and then we apply the result above.

Au fond what’s going on here is the following. We have a theorem that says that if f_1 is a slick function from a nice poset P into itself then P has a nice subposet (or sub-nice-poset) P_1 of fixed points for f_1 . Now let f_2 be a second slick function $P \rightarrow P$. Whack P_1 with f_2 to obtain a nice subposet P_2 (fixed points for f_2) of P_1 . Keep on trucking. To make it work it turns out that you need the f s to commute

Can you make it work if the f s form an ω -sequence, so there are countably many of them. . . ?

(viii)

$\mathbb{N} \rightarrow \mathbb{N}$ is the set of partial functions from \mathbb{N} to \mathbb{N} , thought of as sets of ordered pairs and partially ordered by \subseteq .

Is it complete? Directed-complete? Separative? Which fixed point theorems are applicable?

For each of the following functions $\Phi : (\mathbb{N} \rightarrow \mathbb{N}) \rightarrow (\mathbb{N} \rightarrow \mathbb{N})$, determine (a) whether Φ is order-preserving, and (b) whether it has a fixed point:

- (i) $\Phi(f)(n) = f(n) + 1$ if $f(n)$ is defined, undefined otherwise.
- (ii) $\Phi(f)(n) = f(n) + 1$ if $f(n)$ is defined, $\Phi(f)(n) = 0$ otherwise.
- (iii) $\Phi(f)(n) = f(n - 1) + 1$ if $f(n - 1)$ is defined, $\Phi(f)(n) = 0$ otherwise.

(ix)

Players I and II alternately pick elements (I plays first) from a set A (repetitions allowed: A does not get used up) thereby jointly constructing an element s of A^ω , the set of ω -sequences from A . Every subset $X \subseteq A^\omega$ defines a game $G(X)$ which is won by player I if $s \in X$ and by II otherwise. Give A the discrete topology and A^ω the product topology.

By considering the poset of partial functions $A^{<\omega} \rightarrow \{I\}$ ($A^{<\omega}$ is the set of finite sequences from A) or otherwise prove that if X is open then one of the two players must have a winning strategy.

Discussion Answer

First thing to clear up is the correct definition of open set in the product topology.

The next thing to do is to think of $A^{<\omega}$ as the set of *positions* in the game... *odd* positions when it is II's turn to move, and *even* positions when it is I's turn to play. Observe that, if I wins at all, he wins after finitely many moves. Accordingly, if s is a position such that all elements of A^ω that kick off with s are wins for I, then we are not interested in any proper end-extensions of it. We just label it with a 'I' (to mean that I has won) and have done with it

At this point there are two ways of joining up the dots.

(i) We can define a labelling function by recursion on the end-extension relation between positions. Yes, this means that the empty position is at the top!

The recursion is:

If s is an even position and it has an end-extension by 1 that is labelled 'I' then label it 'I'

If s is an odd position and every end-extension of it by 1 is labelled 'I' then label it 'I'


Then it becomes an exercise in recursion.

(ii) You can take the hint: "...by considering the poset of partial functions $A^{<\omega} \rightarrow \{I, II\}$...". How so? Well, any labelling can be processed by one step of the recursion outlined in (i), so we have a function from labellings to labellings which is monotone, or inflationary—or something. One way or another we reach a fixed point, which will be the labelling we are trying to define by recursion in (i).

Let's think about this fixed point, this *maximal labelling*. The key question is: "Does it label the empty position?". If it does, then I has a winning strategy, which is to always pick a labelled

node when it is his turn. If not, then II has a strategy which is always to play *unlabelled* nodes when it is her turn. That way she stays alive forever and thereby wins.


This is the **Gale-Stewart theorem**, aka “*Open Determinacy*”. It is far from best possible. Best possible is Borel Determinacy, a theorem of D.A.Martin from the 1970’s. A bit fat juicy transfinite induction.

(x) 

$\mathbb{R} = \langle 0, 1, +, \times, \leq \rangle$ is a field. Consider the product $\mathbb{R}^{\mathbb{N}}$ of countably many copies thereof, with operations defined pointwise. Let \mathcal{U} be an ultrafilter $\subseteq \mathcal{P}(\mathbb{N})$ and consider $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$. Prove that it is a field. Is it archimedean?

(xi)

(i)⁺ How many order-preserving injections $\mathbb{R} \rightarrow \mathbb{R}$ are there?

(ii)  Let $\langle X, \leq_X \rangle$ be a total order with no nontrivial order-preserving injection $X \rightarrow X$. Must X be finite?

Discussion Answer

Without the hint in part (i) this is an Impossible-Imre question (from Part II ST&L in earlier years). The point of the hint is that the number of such injections is precisely 2^{\aleph_0} ... which is the same as the number of reals. This fact enables one to construct a counterexample to part (ii) in the form of a set X of reals obtained by a diagonal construction. One wellorders the set of order-preserving injections $\mathbb{R} \hookrightarrow \mathbb{R}$ in a 2^{\aleph_0} -like wellorder [so of course you have to wellorder the continuum] and at each stage α one puts something into X (or into $\mathbb{R} \setminus X$ —i’ll leave that to you to sort out; you use the α th real to bugger up the α th map) to ensure that [the restriction of] the α th order-preserving map $\mathbb{R} \rightarrow \mathbb{R}$ does not inject X into X . You may need to say something about why every order-preserving map $X \rightarrow X$ is a restriction of an order-preserving map $\mathbb{R} \rightarrow \mathbb{R}$. Of course it’s easy if X is dense, but I forget the details. If you want details have a look at: Sierpinski, W. “Sur les types d’ordres des ensembles linéaires”. *Fundamenta Mathematica* **37** (1950) pp 253–264.

Set Theory and Logic, Michaelmas 2016, Sheet 3: Propositional and Predicate Logic

(i)

Show how \wedge , \vee and \neg can each be defined in terms of \rightarrow and \perp . Why can you not define \wedge in terms of \vee ? Can you define \vee in terms of \rightarrow ? Can you define \wedge in terms of \rightarrow and \vee ?

Discussion Answer

When i set these questions i didn’t have particular answers in mind, so i didn’t know what to expect. The following proof that you can’t define ‘ \wedge in terms of ‘ \rightarrow ’ and ‘ \vee ’ is due to Mr. Irving of Churchill.

Think of the four element boolean algebra, with its two extra elements **left** and **right**. Reflect that **left** \rightarrow **right** is **right** and that **right** \rightarrow **left** is **left**. And **left** \vee **right** is of course \top . Consider any complex expression **fake-and**(p, q) with the two letters ' p ' and ' q ' in it, that comically aspires to be conjunction. Consider the valuation that sends p to **left** and sends q to **right**. There is no way it can send **fake-and**(p, q) to \perp , but that's what it would have to do if **fake-and**(p, q) really were $p \wedge q$.

(ii)

- (a) Show that for every countable set A of propositions there is an independent set B of propositions with the same deductive consequences.
- (b) If A is finite show that we can find such a B with $B \subseteq A$.
- (c) Give an example to show that we should not expect $B \subseteq A$ if A is infinite.
- (d) Show that if A is an infinite independent set of propositions then there is no finite set with the same deductive consequences.

(iii)

Explain briefly the relation between truth-tables and Disjunctive Normal Form.

Explain briefly why every propositional formula is equivalent both to a formula in CNF and to a formula in DNF.

Establish that the class of all propositional tautologies is the maximal propositional logic in the sense that any superset of it that is a propositional logic (closed under \models and substitution) is trivial (contains all well-formed formulæ).

(iv)

*A formula (of first-order Logic) is in **Prenex Normal Form** if the quantifiers have been “pulled to the front”—every propositional connective and every atomic subformula is within the scope of every quantifier.*

Explain briefly why every first-order formula is equivalent to one in PNF.

Axiomatise the theory of groups in a signature with '=' and a single three-place relation " x times y is z ". Put your axioms into PNF. What are the quantifier prefixes?

Find a signature for Group Theory which ensures that every substructure of a group is a semigroup-with-1.

(v)

Show that the theory of equality plus one wellfounded relation is not axiomatisable.

(vi)

Write down axioms for a first-order theory T with equality plus a single one-place function symbol f that says that f is bijective and that for no n and no x do we have $f^n(x) = x$.

(a) Is T finitely axiomatisable?

(b) How many countable models does T have (up to isomorphism)?

(c) How many models of cardinality of the continuum does it have (up to isomorphism)?
(You may assume that the continuum is not the union of fewer than 2^{\aleph_0} countable sets, a fact whose proof—were you to attempt it—would need AC.)

(d) Let κ be an uncountable aleph. How many models does T have of size κ ?


(e) Is T complete?

Discussion Answer

For part (d) there are two proofs. One uses the fact that a set of size κ is not the union of fewer than κ countable sets, and the other uses back-and-forth. Neither proof uses AC. Both proofs require accurate attention to detail so supervisors should probably be happy with a hand-wavy proof as long as it shows a satisfactory level of comprehension.

(vii)

Show that monadic predicate logic (one place predicate letters only, without equality and no function symbols) is decidable.

(viii) 

- (a)⁺ Suppose A is a propositional formula and ' p ' is a letter appearing in A .
Explain how to find formulæ A_1 and A_2 not containing ' p ' such that A is logically equivalent to $(A_1 \wedge p) \vee (A_2 \wedge \neg p)$.
- (b) Hence or otherwise establish that, for any two propositional formulæ A and B with $A \models B$, there is a formula C , containing only those propositional letters common to both A and B , such that $A \models C$ and $C \models B$. (Hint: for the base case of the induction on the size of the common vocabulary you will need to think about expressions over the empty vocabulary).

(ix)

Why does T not follow from K and S ?

Show that Peirce's Law: $((A \rightarrow B) \rightarrow A) \rightarrow A$ cannot be deduced from K and S .

(x⁺)

Look up 'monophyletic'. Using only the auxiliary relation "is descended from" give a definition in first-order logic of what it is for a set of lifeforms to be monophyletic.

Discussion Answer

If we write “is descended from” with a “ \succeq ” we can say that P is monophyletic iff $(\exists x)(\forall y)(P(y) \longleftrightarrow x \leq y)$. It a simple real-life application.

(xi)

Is

$$(\forall x)(\exists y)(F(x, y)) \rightarrow (\forall x)(\exists y)(\forall x')(\exists y')[F(x, y) \wedge F(x', y') \wedge (x = x' \rightarrow y = y')]$$


valid?

(xii)

- (a) Show that the theory of fields of characteristic zero is (first-order) axiomatisable but not finitely axiomatisable. Show that the theory of fields of finite characteristic is not first-order axiomatisable.
- (b) Recall that a simple group is one with no nontrivial normal subgroup. Is the theory of simple groups first order?
- (c) A local ring is a ring with a unique maximal ideal. Is the theory of local rings first-order? [Hint: what might the unique maximal ideal be?]
- (d) Is the theory of posets in which every element belongs to a unique maximal antichain first-order?
- (e) A theory T is **equational** iff every axiom of T is of the form $(\forall \vec{x})\Phi$ where Φ is a conjunction of equations between T -terms.

Prove that, if T is equational, then a pointwise product of models of T is another model of T , and substructures and homomorphic images of models of T are models of T .

Which of the theories in (a)–(d) are equational?

(xiii) 

A **type** in a propositional language \mathcal{L} is a countably infinite set of formulæ.

For T an \mathcal{L} -theory a **T -valuation** is an \mathcal{L} -valuation that satisfies T . A valuation v **realises** a type Σ if v satisfies every $\sigma \in \Sigma$. Otherwise v **omits** Σ . We say a theory T **locally omits** a type Σ if, whenever ϕ is a formula such that T proves $\phi \rightarrow \sigma$ for every $\sigma \in \Sigma$, then $T \vdash \neg\phi$.

(a) Prove the following:

Let T be a propositional theory, and $\Sigma \subseteq \mathcal{L}(T)$ a type. If T locally omits Σ then there is a T -valuation omitting Σ .

(b) Prove the following

Let T be a propositional theory and, for each $i \in \mathbb{N}$, let $\Sigma_i \subseteq \mathcal{L}(T)$ be a type. If T locally omits every Σ_i then there is a T -valuation omitting all of the Σ_i .

(xiv)

Prove that, for every formula ϕ in CNF, there is a formula ϕ' which

(i) is satisfiable iff ϕ is;

(ii) is in CNF where every conjunct contains at most three disjuncts.

(Hint: there is no assumption that $\mathcal{L}(\phi') = \mathcal{L}(\phi)$.)

Set Theory and Logic, Michaelmas 2016, Sheet 4: More Predicate Logic and Some Set Theory

(i)⁺

Show that if x is a transitive set, then so are $\bigcup x$ and $\mathcal{P}(x)$. Are the converses true?

(ii)⁺

Show that the Pair-set axiom is deducible from the axioms of empty set, power set, and replacement.

(iii)⁺

Show that $\{z : \neg(\exists u_1, \dots, u_n)((z \in u_1) \wedge (u_1 \in u_2) \wedge \dots \wedge (u_n \in z))\}$ is not a set for any n . What assumptions have you made?

(iv)

Write down sentences in the language of set theory to express the assertions that, for any two sets x and y , the product $x \times y$ and the set y^x of all functions from x to y exist. You may assume that your pairs are Wiener-Kuratowski.

Which axioms of set theory are you going to have to assume if these assertions are to be provable?

Discussion Answer

If you use Wiener-Kuratowski pairs then $\langle x, y \rangle = \{\{x\}, \{x, y\}\}$ and is a subset of $\mathcal{P}^2(\{x, y\})$. Similarly $x \times y$ is a subset of the power set a couple of times of $x \cup y = \bigcup \{x, y\}$. Clearly the set of functions from x to y can be obtained in the same way.

What if you want to establish that these things are sets without knowing what your pairing function is? Imagine the following situation: i want $X \times Y$ and i know that there is a set-theoretic construct $\langle x, y \rangle$, tho' i don't know what it is and i'm not allowed to assume anything other than that it is there and is available. We do the following: fix $y \in Y$ and consider the function class that sends x to $\langle x, y \rangle$. The image of X in this function exists by replacement and it is of course $X \times \{y\}$. So $X \times \{y\}$ exists for all y . Now consider the function class that sends y to $X \times \{y\}$. The image of Y in this function exists by replacement and is $X \times Y$.

So: if we have replacement we can prove that $X \times Y$ exists whatever implementation of pairing-with-unpairing we use. You might like to prove the converse: if $X \times Y$ always exists for all implementations of pairing-with-unpairing then replacement follows.

(v)

- (a) Prove that every normal function $On \rightarrow On$ has a fixed point.
- (b) Prove that the function enumerating the fixed points of a normal function $On \rightarrow On$ is itself normal.
- (c) If α is a regular ordinal and f is a normal function show that f has a fixed point of cofinality α .
- (d) Are any of your fixed points regular?

Discussion Answer

For (a) iterate ω times and take the sup. The reason why this was not on sheet 1 is that one needs replacement if the collection of iterates is to be a set. For (b), let f be your normal function; the new function you need is declared by $g(\alpha + 1) = \sup \{f^n(g(\alpha) + 1) : n < \omega\}$, taking sups at limits. This trick of “add one and keep on trucking” enables you to manufacture strings of fixed points of length as long as you please; pause at any limit and take a sup. This deals with (c). The answer to (d) is that this procedure will never give you a regular fixed point. Not at all clear how we might prove that every normal function has a regular fixed point, and it turns out that this is a very strong assumption—stronger by far than the consistency of ZF.

(vi)

Show that the axiom of choice follows from the assumption that cardinals are totally ordered by \leq_{card} .

(vii)

Explain briefly the equivalence of the four versions of the axiom of foundation given in lectures: (i) The axiom scheme of \in -induction; (ii) The assertion that every set is wellfounded; (iii) Axiom of Regularity; (iv) Every set belongs to the cumulative hierarchy.

(viii)

f is an \in -automorphism if f is a permutation of V that preserves \in : $x \in y \iff f(x) \in f(y)$. Show that a model of ZF (with foundation of course) can have no nontrivial \in -automorphisms.

Give an example to show that the surjectivity condition on f is necessary; that is to say, there are non-trivial injective \in -homomorphisms.

Discussion Answer

The first bit yields to extensionality plus \in -induction. For the second part consider the function f which we define by \in -recursion that sends \emptyset to $\{\emptyset\}$ and thereafter sends x to $f''x$. Possibly worth pointing out to interested students that this establishes the independence of the axiom of extensionality.

(ix)

For the Wiener-Kuratowski ordered pair $\rho(\langle x, y \rangle) = \max(\rho(x), \rho(y)) + 2$. (ρ is set-theoretic rank.)

(a) Can you define a ordered pair such that $\rho(\langle x, y \rangle) = \max(\rho(x), \rho(y)) - 1$?

(b) Can you define a ordered pair such that $\rho(\langle x, y \rangle) = \max(\rho(x), \rho(y)) + 1$?

(c)* Can you define a ordered pair such that $\rho(\langle x, y \rangle) = \max(\rho(x), \rho(y))$ for all
but finitely many x and y ?

(x)

There are various ways of constructing implementations (as sets) of \mathbb{Q} , \mathbb{Z} , \mathbb{R} and \mathbb{C} from an implementation (as sets) of the naturals. For one of these constructions compute the ranks of the sets that have the rôles of \mathbb{Q} , \mathbb{Z} , \mathbb{R} and \mathbb{C} .

Different implementations will almost certainly give you different answers. Are there any lower or upper bounds on the answers you might get?

(xi)

Consider the binary relation E on \mathbb{N} defined by: $n E m$ iff the n th bit (counting from the right, starting at 0) in the binary expansion of m is 1. What can you say about the structure $\langle \mathbb{N}, E \rangle$?

(xii)

Prove that, for each $n \in \mathbb{N}$, there is a set of size \aleph_n . Is there a set of size \aleph_ω ?

(xiii)

Assume that the cartesian product $x \times y$ always exists however you implement ordered pairs. Infer the axiom scheme of replacement.

(xiv)

Assume that every normal function $On \rightarrow On$ has a regular fixed point. Consider the function that enumerates the initial ordinals and deduce that there is a “weak inaccessible” κ . Which axioms of ZF hold in V_κ ?

(xv)

Suppose $\{A_i : i \in I\}$ and $\{B_i : i \in I\}$ are families of sets such that for no $i \in I$ is there is a surjection $A_i \twoheadrightarrow B_i$. Show that there is no surjection $\bigcup_{i \in I} A_i \twoheadrightarrow \prod_{i \in I} B_i$.

You will need the axiom of choice. Is there a converse?

Using these ideas you can show that $\aleph_\omega \neq 2^{\aleph_0}$ without using AC.

Chapter 4

Professor Johnstone's Example Sheets for 2015/6

Thanks (and a bottle of port) are due to Qiaochu Yuan and to Leo Lai (Senior Wrangler in 2016) for some \LaTeX source code; Yanitsa Pehova taught me how to use `geogebra` which i used to do—for example—the pictures for Sheet 1 question 1, and she, too, got a bottle of port. It might even have been two. If you feel like donating \LaTeX code for nice answers you will be laying up treasure in heaven and possibly a bottle of port on earth.

At various points I allude to material in my book *Logic, Induction and Sets* which grew out of the notes from which I lectured a precursor of this course in the last millenium. There should be a copy in your college library. (The colleges have a sweetheart deal with CUP under which they get CUP books cheap . . . So do you lot, come to think of it, so get out there and buy it!)

4.0.1 Sheet 1

Question 1

Write down all possible Hasse diagrams for a poset with four elements. [There are 16 of them.] How many of them are complete?

Discussion

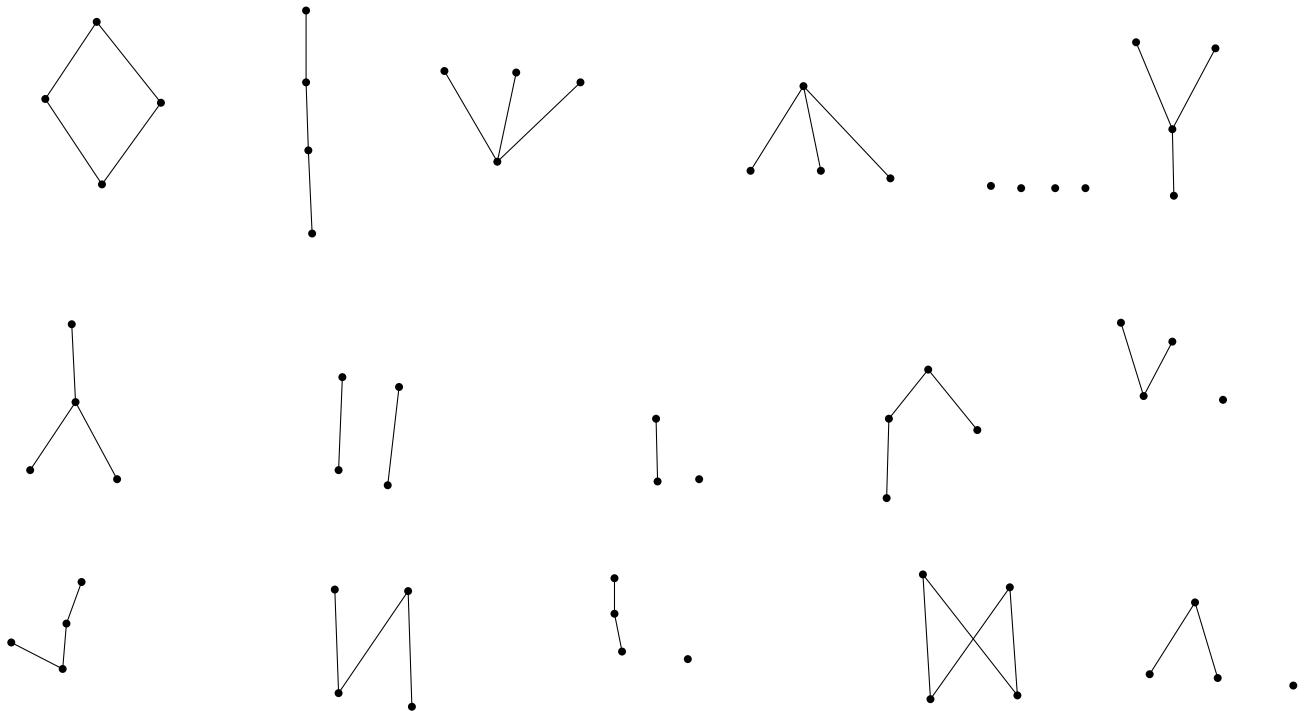
Is there a systematic way of generating them? No easy way, that's for sure. (The set of Hasse diagrams for posets with four elements is a quotient of the set of posets with four elements and in general the cardinalities of quotients are hard to compute.) I find myself wondering how many isomorphism types of partial orders there are on a set of n elements. Not exactly of course, for i see no prospect for an exact formula, but it would be nice to know whether or not there is an exponential lower bound, or a polynomial upper bound. There are $2^{\binom{n}{2}}$ reflexive relations on n elements. How many of them are transitive? My guess is: exponentially many, but i have no exact figure. Let me see. . . How many transitive relations on $n + 1$ things does a given transitive relation on n things extend to? There are $2n$ places where one might put in an edge. The only constraints arise as follows.

- (i) Suppose there is no edge from x to y , where x and y are of the original party of n .

Then we cannot have a new edge from x to a (a is our new chap) as well as a new edge from a to y .

(ii) If there is an edge from x to y and we add a new edge from y to a then we have to add a new edge from x to a .

Anyway, there are 16 isomorphism classes of posets on 4 elements, 2 of which are complete (the two with both a top and a bottom element, co's the empty subset has to have a sup!)



One of my students distinguished



... which are two embeddings of the same poset into the plane. This makes the same interesting point that my Pittsburgh colleague Ken Manders likes to make. When you formalise (= represent something concretely, or *concretise*) you add extra structure and this structure may be spurious. However I don't think this was the point that the question setters were trying to make... apparently the *real* reason for this question is that you weren't taught about Hasse diagrams in 1a. What is the world coming to??

There is a general question here: *How do i know when i've got them all?* This particular instance (before us) of this *general* question isn't so hard that we are prompted to think much about the general question, but a bit of thought won't go amiss. The answer of course is that you have to find a fairly robust way of thinking of these things as mathematical objects and then find a way of classifying them. In this case the obvious thing to do is to identify them with their Hasse diagrams and then classify them—perhaps—in terms of the number of edges they have. But the question still lurks in the shadows: “How can i give a *mathematical* proof that i have got all of them?”

Question 2

Which of the following posets are complete? And which are chain-complete?

(i) *The set of all finite subsets of an arbitrary set, ordered by inclusion.*

It's not complete or even chain complete: The chain $\{\{1\}, \{3\}, \{5\}, \dots\}$ of finite subsets of \mathbb{N} does not have a supremum.

(ii) *The set of cofinite subsets (that is, subsets with finite complements) of an arbitrary set, ordered by inclusion.*

Unless the set is finite the empty set of cofinite subsets lacks a supremum. It would have to be the empty set, and the empty set is a cofinite subset of a set x iff x is finite.

(iii) *The set of all transitive relations on an arbitrary set A , regarded as subsets of $A \times A$ and ordered by inclusion.*

(iv) *The set of all partial orderings of subsets of A , ordered by inclusion.*

Discussion

Readers may have noticed that a union of a hatful of transitive relations might not be a transitive relation, and some of them think that this means that the set of transitive relations on a fixed set is not complete. This is a mistake, beco's it stops too early. You might have to add ordered pairs. And you might have to do it infinitely often. If you let R be $\{\langle 2n, 2n+1 \rangle \in \mathbb{N}\}$ and let S be $\{\langle 2n+1, 2n+2 \rangle \in \mathbb{N}\}$ then R and S are both transitive (vacuously!) but $R \cup S$ isn't, and you have to go on adding ordered pairs to it infinitely often. $R \cup S$ is the successor function on \mathbb{N} , and its transitive closure is $<_{\mathbb{N}}$. “Transitive closure”? *Being a transitive relation is a closedness property and therefore admits a closure operation.*

Observe the failure of symmetry between sup and inf here. There is a connection with failure of distributivity which we will discuss later. Being a subspace of a vector space is a closedness property too [linear combinations] which is connected with the failure of distributivity that we see in Q10 (i).

[PTJ says: *I adopted the nonstandard convention that a chain is a nonempty totally ordered subset of a poset; thus a chain-complete poset does not have to have a least element, and in particular example (ii) is chain-complete.*]

Why do you allow empty subsets (in Q1 every subset has to have a supremum for the poset to be complete) but not empty chains? I think the answer is that chains are *substructures* of posets and substructures are not allowed to be empty ... beco's *structures* are not allowed to be empty. There's a complex debate about this which i don't properly understand and which in any case we don't have to worry about here.

(iii)

One student of mine provided the following blindingly cute construction of the sup of a set S of transitive relations on a given set A . He says

Let $\mathcal{S} = \{\langle x, y \rangle : (\exists z)(\exists s_1, s_2 \in S)(\langle x, z \rangle \in s_1 \wedge \langle z, y \rangle \in s_2)\}$

Observe that \mathcal{S} is a superset of $\bigcup S$. The first challenge is to show that \mathcal{S} is transitive. So suppose $\langle x, y \rangle$ and $\langle y, z \rangle$ are both in \mathcal{S} . If $\langle x, y \rangle \in \mathcal{S}$ this is beco's there is $s_1 \in S$ with $\langle x, y \rangle \in s_1$ and similarly there is $s_2 \in S$ with $\langle y, z \rangle \in s_2$. (These two beco's $\mathcal{S} \supseteq \bigcup S$.) But then $\langle x, z \rangle \in \mathcal{S}$ by definition of \mathcal{S} .

Now we have to show that this \mathcal{S} is the *least* upper bound. But this is easy: if R is a transitive relation that extends all $s \in S$ then, for s_1, s_2 in S , and for all $\langle x, z \rangle \in s_1$ and $\langle z, y \rangle \in s_2$ we must have $\langle x, y \rangle \in R$. But this is simply to say that $R \supseteq \mathcal{S}$.

(iv)

There is a trap here for the unwary. You will have noticed—perhaps even have been told—that in a poset all sups exist iff all infs exist. You can use this to show that the poset of transitive relations on a fixed set is complete, and if you do you have done it without having to think about transitive closures. However if you try to apply the same thinking to (iv) you will end up in the cactus. Ladies and gentlemen it is obvious, is it not, that an arbitrary intersection of [graphs of] partial orderings is a partial ordering? So all infs exist; so all sups exist—the poset of partial orderings on a fixed set is complete! Except it obviously isn't. What has gone wrong?? The point is that an arbitrary intersection of [graphs of] partial orderings is not reliably a partial ordering. . . . What happens if the family is empty?

Question 3

Let P and Q be posets. There are (at least) two possible ways of defining a partial ordering on $P \times Q$: the *pointwise order* is defined by $(a, c) \leq (b, d)$ if and only if $a \leq b$ and $c \leq d$, and the *lexicographic order* by $(a, c) \leq (b, d)$ if and only if either $a < b$ or $(a = b \text{ and } c \leq d)$. Verify that both of these are partial orders. For each of the following properties, determine whether $P \times Q$ has the property (a) in the pointwise ordering, and (b) in the lexicographic ordering, if both P and Q have the property:

- (i) being complete;
- (ii) being totally ordered;
- (iii) being chain-complete.

Discussion

Observe that altho' $(0, 1] \times (0, 1]$ is the lexicographic product of two chain-complete posets, it is nevertheless not chain-complete: the chain $\{\langle x, 1/2 \rangle : 0 < x < 1\}$ has no least upper bound. Incidentally this shows (in case you were wondering, as i had been) that this lexicographic product isn't iso to \mathbb{R} (never mind the points at infinity) beco's \mathbb{R} is complete!

Incidentally beware of the double use of the word 'complete'. When people say that \mathbb{R} is complete they don't mean that every set of reals has a lub; they mean only that every *bounded* set of reals has a lub!

Question 4

Discussion

The nicest and most natural example of an order-reversing map with no fixed point is complementation in a boolean algebra.

For the second part, if f is an order-reversing function from a complete poset into itself then f^2 is order preserving and has a fixed point.

For the rider, let a be the least fixed point for f^2 . $a = f^2(a)$ so by injectivity of f we have $f(a) = f(f^2(a)) = f^2(f(a))$ so $f(a)$, too, is a fixed point for f^2 . But a was the *least* fixed point, so we must have $a \leq f(a)$.

Why on earth would you be looking for an order-reversing function to have a fixed point? More often than you might think. (And I don't just mean trivial cases like $1/2$ is a fixed point for the order-reversing function $x \mapsto (1 - x)$.) If you think a *species* in Biology is defined in terms of “can mate to produce viable offspring” you rapidly discover a characterisation in terms of fixed points for an order-reversing function. Have a look, too at this old tripos question (It was 2002:B2:11b).

Provide back-
ence

1. State Zorn's lemma.
2. Let U be an arbitrary set and $\mathcal{P}(U)$ be the power set of U . For X a subset of $\mathcal{P}(U)$, the **dual** X^\vee of X is the set $\{y \subseteq U : (\forall x \in X)(y \cap x \neq \emptyset)\}$.
3. Is the function $X \mapsto X^\vee$ monotone? Comment.
4. By considering the poset of those subsets of $\mathcal{P}(X)$ that are subsets of their duals, or otherwise, show that there are $X = X^\vee$.
5. What can you say about the fixed points of $X \mapsto X^\vee$ on the assumption that U is finite?

Question 5

For each of the following functions $\Phi : (\mathbb{N} \rightarrow \mathbb{N}) \rightarrow (\mathbb{N} \rightarrow \mathbb{N})$, determine (a) whether Φ is order-preserving, and (b) whether it has a fixed point:

- (i) $\Phi(f)(n) = f(n) + 1$ if $f(n)$ is defined, undefined otherwise.
- (ii) $\Phi(f)(n) = f(n) + 1$ if $f(n)$ is defined, $\Phi(f)(n) = 0$ otherwise.
- (iii) $\Phi(f)(n) = f(n - 1) + 1$ if $f(n - 1)$ is defined, $\Phi(f)(n) = 0$ otherwise.

[PTJ says: (i) is order-preserving, but its only fixed point is the empty function. (iii) is not order-preserving, but has a unique (and obvious) fixed point.]

Discussion

Values of the last two operations are always total functions so the operations can't be order-preserving!

Part (iii) requires care. Suppose $f = \Phi(f)$; what is $\Phi(f)(0)$? To ascertain the value of $\Phi(f)(0)$ we have to look at $f(0 - 1)$...and that will crash, so $\Phi(f)(0)$ traps the failure and returns 0. Thereafter the recursion proceeds smoothly to give us the identity function.

Question 6

(Tripos IIA 98107). Let P be a chain-complete poset with a least element, and $f : P \rightarrow P$ an order-preserving map. Show that the set of fixed points of f has a least element and is chain-complete in the ordering it inherits from P . Deduce that if f_1, f_2, \dots, f_n are order-preserving maps $P \rightarrow P$ which commute with each other (i.e. $f_i \circ f_j = f_j \circ f_i$ for all i, j), then they have a common fixed point. Show by an example that two order-preserving maps $P \rightarrow P$ which do not commute with each other need not have a common fixed point.

Discussion

[PTJ says: The first part is similar to part (ii) of question 4. For the second, first show that each of f_2, \dots, f_n maps the set of fixed points of f_1 into itself, and then use induction on n . (It's worth pointing out the similarity with the technique—which they should have seen before—for finding a common eigenvector of a family of commuting endomorphisms.) The simplest counterexample for the last part is to take two distinct constant maps].

If $f_1 \dots f_n$ are order-preserving maps $P \rightarrow P$ which commute with each other then their set of common fixed points is chain-complete and has a least element.

Proof: by induction on n . Base case easy! If $n > 1$ by induction hypothesis the set F of common fixed points of $f_1 \dots f_{n-1}$ is chain complete and has a least element. Since f_n commutes with all f_k with $k < n$, f_n sends F into itself and then we apply the result above.

Au fond what's going on here is the following. We have a theorem that says that if f_1 is a slick function from a nice poset P into itself then P has a nice subposet (or sub-nice-poset) P_1 of fixed points for f_1 . Now let f_2 be a second slick function $P \rightarrow P$. Whack P_1 with f_2 to obtain a nice subposet P_2 (fixed points for f_2) of P_1 . Keep on trucking. To make it work it turns out that you need the f s to commute

Can you make it work if the f s form an ω -sequence, so there are countably many of them...?

Question 7

Discussion

A poset P is **Bourbakian** iff every order-preserving map $f : P \rightarrow P$ has a least fixed point. I've never heard this boldfaced word used anywhere else so i think it's a PTJ-ism. No, he's not Armenian; there is a **rather** cute proof of Bourbaki-Witt by someone from those parts [which i can show you] but he's a Georgian.

And don't go away with the idea that the definition is supposed to capture those posets that obey the theorem of Bourbaki-Witt; B-W states that in a chain-complete poset every **inflationary** function has a fixed point. Here we are concerned with **order-preserving** functions.

The aim of this question is to show that the class of Bourbakian posets is closed under substructure and products, which means it's an algebraic variety or something.

I'm going to restate the question by numbering its parts, to make the discussion answer easier to follow.

- (i) Show that if P is Bourbakian and $p \in P$ then the substructure $\downarrow p = \{p' \in P : p' \leq p\}$ of P consisting of things $\leq p$ is Bourbakian.

- (ii) Now suppose $f : P \rightarrow P$ and $g : P \rightarrow P$ both order preserving, and $(\forall x \in P)(f(x) \leq g(x))$. Show that $\mu(f) \leq \mu(g)$.
- (iii) So let P and Q be Bourbakian, and let $h : P \times Q \rightarrow P \times Q$ be order-preserving (wrt the pointwise product ordering). Now $h(x, y)$ is of course an ordered pair, so we can unpair it to get two functions h_1 and h_2 s.t. $(\forall x \in P)(\forall y \in Q)(h(x, y) = \langle h_1(x, y), h_2(x, y) \rangle)$.
Fix any $x_0 \in P$ and prove that the function $g_{x_0} : Q \rightarrow Q$ defined by $g_{x_0}(y) =: h_2(x_0, y)$ is order-preserving.
- (iv) Show that $x \mapsto \mu(g_x) : P \rightarrow Q$ is order-preserving.
- (v) Show that $\langle \mu(f), \mu(g_{\mu(f)}) \rangle$ is the least fixed point of h .

[This discussion answer is still work-in-progress]

- (i) First we show that if P is Bourbakian and $p \in P$ the substructure $\downarrow p = \{p' \in P : p' \leq p\}$ of P consisting of things $\leq p$ is likewise Bourbakian. Actually he's shown it for us, nice man. Sse $f : \downarrow p \rightarrow \downarrow p$ is order-preserving. Extend it to a function (also notated ' f ') $f : P \rightarrow P$ by ordaining that $f(p') = p$ whenever $p' \not\leq p$. By assumption that P is Bourbakian we conclude that f has a least fixed point, called $\mu(f)$. We want $\mu(f) \leq p$. But this is easy—all values of f are $\leq p$. And let's hang onto this ' μ ' notation for the least fixed point for future use.
- (ii) Now suppose $f : P \rightarrow P$ and $g : P \rightarrow P$ both order preserving, and $(\forall x \in P)(f(x) \leq g(x))$. Consider $f \upharpoonright \mu(g)$. By the above, $\downarrow (\mu(g))$ is Bourbakian. Suppose $x \in \downarrow (\mu(g))$. We have $f(x) \leq f(\mu(g)) \leq g(\mu(g)) = \mu(g)$, so we have $f \upharpoonright \mu(g) : \downarrow \mu(g) \rightarrow \downarrow \mu(g)$, so (since $\downarrow \mu(g)$ is Bourbakian) $f \upharpoonright \mu(g)$ has a least fixed point, which is clearly bounded above by $\mu(g)$ as desired.
- (iii) Fix $x \in P$. $g_x(y) =: h_2(x, y)$, order-preserving: if $y \leq y'$ then $\langle x, y \rangle \leq \langle x, y' \rangle$ in the pointwise order so $h(x, y) \leq h(x, y')$ (again, in the pointwise order) and $\langle h_1(x, y), h_2(x, y) \rangle \leq \langle h_1(x, y'), h_2(x, y') \rangle$. In particular $h_1(x, y) \leq h_1(x, y')$ which is to say $g_x(y) \leq g_x(y')$. So g_x is order-preserving as desired.
- (iv) [show that $x \mapsto \mu(g_x) : P \rightarrow Q$ is order-preserving]
Suppose we have $x \leq x' \in P$. To prove that $\mu(g_x) \leq \mu(g_{x'})$ it will be sufficient (by our answer to (iii)) to show that, for all $y \in Q$, $g_x(y) \leq g_{x'}(y)$. Now, for all y in Q , we have $\langle x, y \rangle \leq \langle x', y \rangle$ whence $h(x, y) \leq h(x', y)$ which is to say (unroll dfn of h_1 and h_2) $\langle h_1(x, y), h_2(x, y) \rangle \leq \langle h_1(x', y), h_2(x', y) \rangle$ giving $g_x(y) \leq g_{x'}(y)$.
- (v) Now define $f(x) =: h_1(x, \mu(g_x))$. Suppose $x \leq x' \in P$. Then $\langle x, \mu(g_x) \rangle \leq \langle x', \mu(g_{x'}) \rangle$ (beco's $\mu(g_x) \leq \mu(g_{x'})$ as above, part ??). So $h_1(x, \mu(g_x)) \leq h_1(x', \mu(g_{x'}))$ ($h_1 : P \times Q \rightarrow P$ is order preserving, by above part ??)

Now to show that a product of Bourbakian posets is Bourbakian. So let P and Q be Bourbakian, and let $h : P \times Q \rightarrow P \times Q$ be order-preserving (wrt the pointwise product ordering). Now $h(x, y)$ is of course an ordered pair, so we can unpair it to get two functions h_1 and h_2 s.t. $(\forall x \in P)(\forall y \in Q)(h(x, y) = \langle h_1(x, y), h_2(x, y) \rangle)$. Observe [check this] that $h_1 : P \times Q \rightarrow P$ and $h_2 : P \times Q \rightarrow Q$ are both order-preserving.

Fix any $x_0 \in P$ and consider the map $g_{x_0} : Q \rightarrow Q$ defined by $g_{x_0}(y) =: h_2(x, y)$. We need to show that it is order-preserving. That is to say, we want $(\forall y \leq_Q y' \in Q)(g_{x_0}(y) \leq g_{x_0}(y'))$ which is to say $(\forall y \leq_Q y' \in Q)(h_2(x_0, y) \leq h_2(x_0, y'))$ which is immediate beco's h_2 is order-preserving from $P \times Q \rightarrow Q$.

In his message to supervisors Prof Johnstone says “students could be asked to think about whether or not Bourbakian posets are neccessarily chain-complete (the answer is yes if you assume AC ...)”

Why might a Bourbakian poset be chain-complete? Obviously we have to use its Bourbakian nature to add sups to all chains. Let $P = \langle P, \leq_P \rangle$ be Bourbakian and let $C \subseteq P$ be a chain.

I need to find a cute proof of this! Apparently it's hard, and i haven't got it yet.

(Tripos IIA 95408, modified).

Let P and Q be chain-complete posets with least elements, and let $h : (P \times Q) \rightarrow P \times Q$ be a map which is order-preserving with respect to the pointwise ordering (cf. question 3). We denote the two components of the ordered pair $h(x, y)$ by $h_1(x, y)$ and $h_2(x, y)$ respectively. (i) Show that, for each fixed $x \in P$, the mapping $g_x : Q \rightarrow Q$ defined by $g_x(y) = h_2(x, y)$ is order-preserving. Let $m(x)$ denote its least fixed point. (ii) Show that the map $f : P \rightarrow P$ defined by $f(x) = h_1(x, m(x))$ is order-preserving. Let x_0 denote its least fixed point. (iii) Show that $(x_0, m(x_0))$ is the least fixed point of h .

Discussion

[PTJ says: *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 $x_1 \leq x_2$ then $m(x_2)$ is a ‘post-fixed point’ of g_{x_1} (that is, $m(x_2) \geq g_{x_1}(m(x_2))$), and so $\{y \in Q : y \leq m(x_2)\}$ is a ‘closed set’ in the sense used in the construction of the least fixed point $m(x_1)$ of g_{x_1} .]*

IIA 95408

My answer to the original..

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

(ii) m is order-preserving.

Proof: Suppose $x_1 \leq x_2 \in P$. Set $Y = \{y \in Q : y \leq m(x_2)\}$. For $y \in Y$ we have $g_{x_1}(y) = h_2(\langle x_1, y \rangle) \leq h_2(\langle x_2, y \rangle) = m(x_2)$ 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 $g_{x_1} \upharpoonright Y$ has a fixed point $y_0 \in Y$ with $m(x_1) \leq y_0 \leq m(x_2)$ and m is order-preserving.

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

(iii) $h(x_0, m(x_0)) = \langle f(x_0), g_{x_0}(m(x_0)) \rangle = \langle x_0, m(x_0) \rangle$ 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 $y \geq m(x)$.

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

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

$h(a, b) \leq h(x, m(x)) \leq \langle h_1(x, y), h_2(x, m(x)) \rangle = \langle x, m(x) \rangle$ 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 $\langle x, y \rangle \leq \langle x', y' \rangle \leq \langle x, m(x) \rangle$. But $\langle x, y \rangle \geq \langle x, m(x) \rangle$ so $\langle x, y \rangle = \langle x, m(x) \rangle$.

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

Question 8

A poset (P, \leq) is said to be *inductive* if each chain in P has an upper bound (but not necessarily a least upper bound). The usual statement of Zorn's Lemma says that every inductive poset (and not merely every chain-complete poset) has enough maximal elements.

(i) Give an example of a poset which is inductive but not chain-complete. (The best example of an inductive poset that is not chain complete would be any closed interval in the rationals! PTJ's preferred example is \mathbb{N} with two added points at infinity.)

(ii) If (P, \leq) is any poset, let C denote the set of all chains in P , ordered by inclusion. Show that C is chain-complete.

(iii) If M is a maximal element of the poset C just defined, show that any upper bound for M in P is (a) a member of M , and (b) a maximal element of P .

(iv) Deduce the usual statement of Zorn's Lemma from the version proved in lectures.

Discussion

Of course this question is all about the completion of incomplete objects. You can obtain \mathbb{R} from \mathbb{Q} . The use of chains here is a close analogue of the use of Dedekind cuts or Cauchy sequences. The slight difference here is that whereas when you complete the rationals you have a canonical injection from the original poset into the completion, here you don't, or at least you seem not to. (To which chain-with-top-element do you send a member of the original poset? No point in sending x to $\{x\}$ co's that's not order-preserving!) You can also use the chain-complete poset of directed subsets of the original poset. ['Directed'? A poset is directed iff every finite subset has an upper bound.] If you do that then you can send each element x of the original poset to the directed subset $\{y : y \leq x\}$ so you get an injective homomorphism. Of course *one* of the ways of completing the rationals to obtain the reals is to use Cauchy sequences. That needs a *metric* structure which we haven't got here. If you have another—different—metric structure, you will get a different completion. Look up p -adic numbers They're dead sexy.

Question 9

Use Zorn's Lemma to prove

(i) that every partial ordering \leq_X on a set X can be extended to a total ordering of X ;

(ii) that, for any two sets A and B , there exists either an injection $A \rightarrow B$ or an injection $B \rightarrow A$.

Discussion

(i) Give the set of partial orders on X the containment partial order as subsets of $X \times X$. The resulting partial order is chain-complete, since the union of a nested sequence of partial orders is still a partial order. In fact the containment order is *directed-complete*: (see page 99 for dfn of ‘directed’) every *directed* subset has a l.u.b. By Zorn’s lemma, it follows that, given any partial order \leq on X , there exist maximal partial orders extending it. Let \leq' be one such and let us establish that \leq' is a total order. Suppose it were not, and that we had $a, b \in X$ with $a \not\leq' b$ and $b \not\leq' a$. Then the relation \leq'' defined by $x \leq'' y$ iff $x \leq' y \vee (x \leq' a \wedge b \leq' y)$ is a partial order [easy to check that it is a partial order] properly extending \leq' —contradicting maximality of \leq' . ■

The poset of partial-orderings-under-inclusion is the obvious poset to use in (i), but one can also consider the set of total orders of subsets of X that are compatible with \leq_X —again, ordered by \subseteq . [You can guess what I mean by ‘compatible’ and it will do you no harm to write out a formula that captures it.] There is no particular advantage to using this special partial order in this case, but a modification of it comes in handy in sheet 4.

Actually, thinking about the poset-of-partial-orderings... we have been thinking of the partial orderings as *sets of ordered pairs* ... in which case the order relation is plain old \subseteq , set inclusion. But it is probably worth making the point that you don’t *have* to think of them as sets of ordered pairs. You can just think of them as relations (whatever *they* are!) and say that a partial order \leq is below (however you write that!) another partial order \leq' iff $(\forall xy)(x \leq y \rightarrow x \leq' y)$. You don’t have to coerce everything in sight into being a set if you don’t feel like it. Admittedly the coercion makes for a more uniform treatment but it doesn’t actually shed any light on the proofs.

For (ii) you of course consider the chain-complete poset of partial bijections.

Question 10

Which of the following posets are lattices? Of those that are lattices, which are distributive?

- (i) The set of all subspaces of a vector space, ordered by inclusion.
- (ii) The set of natural numbers, ordered by divisibility.
- (iii) The set Σ^* of all words over an alphabet Σ ...
- (iv) The unit square $[0, 1] \times [0, 1]$ with the lexicographic ordering (cf. question 3).

[PTJ says: (iv) is totally ordered and hence distributive. (v) is a lattice, but not distributive. The others are obvious: Remind them that they did the non-distributivity of (i) on a Linear Maths example sheet.]

Discussion

Every total order is a distributive lattice. (Check all six cases this once; trust me ... you’ll never have to do it again).

Distributivity. There is a representation theorem of some kind that says that every lattice $\mathfrak{X} = \langle X, \leq, \vee, \wedge \rangle$ is isomorphic to a structure $\mathfrak{Y} = \langle \mathcal{Y}, \subseteq, \cup, \cap \rangle$. Now \mathcal{Y} is obviously going to be a subset of $\mathcal{P}(Y)$ for some set Y . If the \cup and \cap of \mathfrak{Y} are the restrictions to \mathcal{Y} of the \cup and \cap of

$\mathcal{P}(Y)$ then \mathfrak{X} is distributive (co's \cup and \cap are). If \cup -in-the-sense of \mathfrak{Y} is the restriction to \mathfrak{Y} of the \cup of $\mathcal{P}(Y)$ but \cap -in-the-sense of \mathfrak{Y} is *not* the restriction to \mathfrak{Y} of the \cap of $\mathcal{P}(Y)$ then all bets are off, and you shouldn't expect distributivity¹. Applying this reflection to (i) you observe that the inf of two subspaces is their intersection, whereas their sup is not their union but the subspace *spanned by* their union. Once you are alerted it is easy to come up with an example to illustrate nondistributivity.

Observe that, nevertheless, the subspaces of a given vector space do form a chain-complete poset.

For (ii) ... How do we prove that \mathbb{N} with divisibility is distributive? (If, that is—like me—you don't want to think about repeated factors!) Here are two proofs.

(I)

For each p , the powers of p are totally ordered by divisibility and are therefore a distributive lattice. If you think a lattice has to have a top element then 0 fits the bill. We will need two important facts which you should check:

- Arbitrary products of distributive lattices are distributive lattices, and
- Sublattices of distributive lattices are distributive lattices.

So we take the product of all these distributive lattices and throw away the infinite elements (except 0 of course). \mathbb{N} -with-divisibility is now a distributive lattice, being a substructure of a product of distributive lattices.

(II)

Define an injection $\mathbb{N} \hookrightarrow \mathcal{P}_{\aleph_0}(\mathbb{N} \times \mathbb{N})$ [that's the set of finite subsets of $\mathbb{N} \times \mathbb{N}$] by sending n to the union of all the sets $\{p\} \times \{m \in \mathbb{N} : 1 \leq m \leq k\}$ where p is a prime and p^k is the largest power of p that divides n , so that (for example) we send 60 to $\{\langle 2, 1 \rangle, \langle 2, 2 \rangle, \langle 3, 1 \rangle, \langle 5, 1 \rangle\}$. HCF gets sent to \cap and LCM gets sent to \cup . We know that \cup and \cap distribute and therefore HCF and LCM must distribute too.

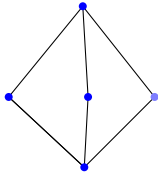
(iii) is obviously not a lattice—as one of my supervisees wrote 'PURE' and 'APPLIED' have no join!

(iv) looks scary but it ain't. The ordering is total, and so is a distributive lattice!

Question 11

Let L be the (five-element) lattice of subgroups of the non-cyclic group of order four. Show that there are no lattice homomorphisms $L \rightarrow 2$.

¹Dishonesty of one of the operations doesn't *guarantee* failure of distributivity, but you have to work quite hard to find an example where it doesn't have this effect. Consider the lattice of regular open subsets of a topological space. \wedge is honest, but \vee is not: $x \vee y$ is the interior of the closure of $x \cup y$ which may be a proper superset of $x \cup y$. Observe that distributivity fails for the infinitary \bigwedge and \bigvee .



[PTJ sez: Each of the three two-element subgroups is complementary to each of the others, so the assumption that a putative homomorphism sends one of them either to 0 or to 1 leads to a contradiction.]

Question 12

A *Boolean ring* is a ring (with 1) in which every element x satisfies $x^2 = x$.

(i) Show that in a Boolean ring R we have $x + x = 0$ and $xy = yx$ for all $x, y \in R$.

(ii) Show that, if we define a relation \leq on R by setting $x \leq y$ if and only if $xy = x$, then \leq is a partial order with respect to which R is a Boolean algebra.

(iii) Show that any Boolean algebra has the structure of a Boolean ring if we define multiplication to be binary meet and addition to be ‘symmetric difference’ (that is, $x + y = (x \wedge \neg y) \vee (\neg x \wedge y)$).

[PTJ says For (i), expand $(x + x)^2$ and then $(x + y)^2$. The rest is straightforward.]

Question 13

A lattice L is called a *Heyting Algebra* if, for any two elements $a, b \in L$, there is a unique largest $c \in L$ with $c \wedge a \leq b$ (this c is usually denoted $a \Rightarrow b$).

(i) Show that every Boolean algebra is a Heyting algebra.

In any Boolean algebra we can define $a \Rightarrow b$ to be $\bar{a} \vee b$.

(ii) Show that every Heyting algebra is a distributive lattice.

An answer from Mr. Pavlovich.

Let c be maximal such that $x \wedge c \leq (x \wedge y) \vee (x \wedge z)$ (i.e. c is $x \Rightarrow ((x \wedge y) \vee (x \wedge z))$.)

Both y and z satisfy this inequality, so $c \geq y$ and $c \geq z$, whence $c \geq y \vee z$.

This gives

$$x \wedge (y \vee z) \leq x \wedge c \leq (x \wedge y) \vee (x \wedge z).$$

Observe that apparently we cannot derive this inequality unless c (which is $x \Rightarrow ((x \wedge y) \vee (x \wedge z))$) exists.

For the inequality going the other way, reason as follows:

$$y \leq (y \vee z) \text{ so } (x \wedge y) \leq x \wedge (y \vee z)$$

$$z \leq (y \vee z) \text{ so } (x \wedge z) \leq x \wedge (y \vee z)$$

so

$$(x \wedge y) \vee (x \wedge z) \leq x \wedge (y \vee z)$$

(iii) Show that a complete lattice L is a Heyting algebra if and only if it satisfies the ‘infinite distributive law’

$$a \wedge \bigvee S = \bigvee \{a \wedge s : s \in S\}$$

for all $a \in L$ and $S \subseteq L$. [Hint: consider $\bigvee \{c \in L : c \wedge a \leq b\}$.] Deduce that every finite distributive lattice is a Heyting algebra.

[still need to prove that a Heyting algebra is a complete lattice satisfying the infinitary distributive law]

For the reverse direction (a complete lattice satisfying the infinite distributive law is Heyting) argue as follows.

Let L be a complete lattice, and $a, b \in L$. Set $c =: \bigvee \{x : x \wedge a \leq b\}$ (which is defined, since L is complete). Then $a \wedge c = \bigvee \{a \wedge x : x \wedge a \leq b\}$ which is of course $\leq b$. Clearly c is the greatest x s.t. $x \wedge a \leq b$. So c is $a \rightarrow b$ and L is a Heyting algebra as desired.

The riders:

Since any finite distributive lattice is complete it follows that every finite distributive lattice is Heyting.

The lattice of open sets of a topological space is a complete poset and satisfies the infinite distributive law. [should find something to say about this... PTJ says that this last bit “is most easily proved using (iii), plus the fact that the finite meet and arbitrary join operations in $\mathcal{O}(X)$ coincide with the set-theoretic intersection and union; but it can also be done directly by defining $U \rightarrow V$ to be the interior of $V \cup (X \setminus U)$. Of course, $\mathcal{O}(X)$ is a Boolean algebra iff every open subset of X is closed; for a T_0 -space X , this is equivalent to X being discrete.”]

4.0.2 Sheet 2

Question 1

This duplicates q1 on sheet 1 of Dr Russell’s set

Which of the following propositional formulæ are tautologies?

- (i) $((p \rightarrow (q \rightarrow r)) \rightarrow (q \rightarrow (p \rightarrow r)))$;
- (ii) $((p \rightarrow q) \rightarrow r) \rightarrow ((q \rightarrow p) \rightarrow r)$;
- (iii) $((p \rightarrow q) \rightarrow p) \rightarrow p$;
- (iv) $((p \rightarrow (p \rightarrow q)) \rightarrow p)$.

Question 2

Use the Deduction Theorem to show that the converse of the third axiom (i.e. the formula $(p \rightarrow \neg \neg p)$) is a theorem of the propositional calculus.

[PTJ says Easiest to use the Deduction Theorem twice, to reduce the problem to proving $\{p, \neg p\} \vdash \perp$.]

Question 3

(Tripos IB 92405). Let t be a propositional formula not involving the constant \perp , and let $t' = t[\perp/p]$ be the formula obtained from t by substituting \perp for all occurrences of a particular propositional variable p in t . Suppose that t' is a tautology but t is not; show that any proof of t' in the propositional calculus must involve an instance of the third axiom. Does this result remain true (a) if t is allowed to contain occurrences of \perp , or (b) if \perp is replaced by \top ?

[PTJ says: *The point is that, if one had a proof of t' using only (K), (S) and (MP), one could convert it into a proof of t by replacing each occurrence of \perp by a new propositional variable. This doesn't work for \top because $\top = (\perp \rightarrow \perp)$ isn't a primitive symbol.*]

...we can prove by induction that if A is derivable from K and S and contains \perp then \perp can be replaced in A —and indeed throughout the proof of A —by some new letter not in A ; the transformed proof is still a proof within the meaning of the act so the modified A is still deducible from K and S . However the result of modifying the third axiom in this way is not a propositional tautology, and therefore cannot be deduced from K and S .

Question 4

Write down a proof of $(\perp \rightarrow q)$ in the propositional calculus [hint: observe the result of question 4 below], and thence write down a deduction of $(p \rightarrow q)$ from $\{\neg p\}$.

Question 5

(Tripos IIA/IIB 98308, modified).

Show that if there is a deduction of t from $S \cup \{s\}$ in n lines (that is, consisting of n consecutive formulae), then $(s \rightarrow t)$ is deducible from S in at most $3n + 2$ lines. Show further that there is a deduction of \perp from $\{((p \rightarrow q) \rightarrow p), (p \rightarrow \perp)\}$ in 16 lines [hint: cf. question 2], and thence calculate an upper bound for the length of a proof of the tautology of question 1(iii).

[PTJ says: *The first estimate comes from examining the proof of the Deduction Theorem: each line in the deduction of t expands to three lines in the deduction of $(s \rightarrow t)$, unless the line consists of the formula s , in which case it expands to five lines (but this needn't occur more than once!). It follows that we can deduce $\neg\neg p$ from $((p \rightarrow q) \rightarrow p)$ in 50 lines; two more lines are needed to convert this into a deduction of p , and then a further application of the Deduction Theorem to obtain a 158-line proof of $((p \rightarrow q) \rightarrow p) \rightarrow p$ (Peirce's Law). This is not the shortest possible proof of Peirce's Law; the best I have been able to do takes 40 lines. Note that any proof of it must involve tertium non datur, since it's not an intuitionistic tautology.*]

What's the point?

By the time you have reached this point, you will have spent a lot of time trying to find proofs of obvious tautologies and tearing your hair out in the process. This isn't a mere piece of torture, it is a Character-Forming Experience². This will cause you to wonder: is there a quick way of finding a proof if there is one? This is actually a mathematically substantial question, even tho' it mightn't look like one. If there is a proof system for propositional logic which gives short proofs to short tautologies then $NP = co-NP$. This is because if there is such a proof system then the set of

²Professor Johnstone is not a mindless sadist; he is a refined and thoughtful sadist who has your interests at heart.

propositional tautologies becomes NP. The set of non-tautologies is known to be NP (obvious: guess a falsifying valuation and check in linear time that it is falsifying) and in fact it is NP-*complete*, which is to say it is as hard as an NP problem can be. If even one NP-complete problem is in co-NP then all of them are, so $\text{NP} = \text{co-NP}$. And—as you may know—the question of whether-or-not $\text{NP} = \text{co-NP}$ is open. (It's not *quite* the same as the question whether or not $\text{P} = \text{NP}$; if $\text{P} = \text{NP}$ then certainly $\text{NP} = \text{co-NP}$, but nobody knows how to prove the converse.)

Question 6

(Tripos IB 89405).

The beliefs of each member i of a finite non-empty set I of individuals are represented by a consistent, deductively closed set S_i of propositional formulæ. Show that the set

$$\{t : \text{all members of } I \text{ believe } t\}$$

is consistent and deductively closed. Is the set

$$\{t : \text{over half the members of } I \text{ believe } t\}$$

deductively closed or consistent?

Question 7

Discussion

One direction is easy: if G is orderable so are all its subgroups, in particular all its finitely generated subgroups. Obvious tho' this is, it's a useful snotty-logician-opportunity to make the point that this happens beco's the theory of orderable groups is \forall^* , so the class of its models is closed under substructure.

For the converse we set up a propositional language and exploit propositional compactness. The language has, for each pair of distinct elements $a, b \in G$, a propositional letter $p_{a,b}$. (Secretly the meaning of $p_{a,b}$ is that $a < b$). The propositional theory to which we are going to apply compactness has the axiom schemes:

$$\begin{aligned} p_{a,b} &\rightarrow p_{ac,bc} \text{ for all } a, b, c \in G; \\ p_{a,b} &\rightarrow p_{ca,cb} \text{ for all } a, b, c \in G; \\ p_{a,b} &\rightarrow (p_{b,c} \rightarrow p_{a,c}) \text{ for all } a, b, c \in G; \\ p_{a,b} &\text{ XOR } p_{b,a} \text{ for all } a, b \in G. \end{aligned}$$

The first two state that the order respects group multiplication, and the third and fourth assert that the order is total.

Any finite set T' of these axioms is consistent beco's each finite subset mentions only finitely many elements of G . For each such T' consider the subgroup $G^{(T')}$ of G generated by the elements mentioned in the subscripts of the propositional letters appearing in T' . This is a finitely generated subgroup of G and is accordingly orderable by hypothesis. Any ordering of $G^{(T')}$ gives a valuation which satisfies T' .

For the second part

It's obvious that any orderable group is torsion-free. For the other direction it will suffice to prove that every finitely-generated torsion-free abelian group is orderable (because then we can use the compactness result we've just seen). But the structure theorem sez that every finitely generated abelian group is a direct product of cyclic groups, and we need only the special case of this for torsion-free abelian groups. In that case the cyclic groups are obviously \mathbb{Z} , which is certainly orderable.

If you want to check your understanding of this process, state and prove an analogous result for *circularly orderable* groups. What?!? Integers mod n have a circular ordering that interacts satisfactorily with the (additive) group structure [*not* the multiplicative structure!]. But the circular-order relation is a *ternary* relation! I have a model tripos question on this, linked from my teaching page:

http://www.dpmms.cam.ac.uk/~tf/cam_only/oldtriposquestions.pdf

A note on notation.... The proof above uses propositional letters in the style $p_{a,b}$. I said “Secretly the meaning of $p_{a,b}$ is that $a < b$ ”. The significance of the word ‘secret’ is that of course the propositional logic cannot “see” the ‘ a ’ and the ‘ b ’. As far as the logic is concerned ‘ $p_{a,b}$ ’ is just a propositional letter: it has no internal structure that the logic can see—the internal structure that we can see is purely typographical.

However, one can make the structure explicit, and conduct the proof instead in a fragment of first-order logic which has predicates and functions, but has no variables—and therefore no quantifiers. That way, instead of having a propositional letter $p_{a,b}$ one has the atomic formula ‘ $a \leq b$ ’. I might write out the details later if i have time, or if i am badgered into it.

The Order Extension Principle

While we are about it here is a treatment of the order extension principle in the same style.

To prove **OEP** by compactness the easy thing to do is to invent a name for every element of the domain of your partial order, and add $a < b$ to the theory of total order for every a and b in $\langle X, R \rangle$ such that $R(a, b)$, obtaining a theory T . We have to do two things (i) Use compactness to prove T consistent and (ii) observe that $\langle X, R \rangle$ can be embedded in any model of T in an order-preserving way. To do (i) we have to prove that any finite fragment of T is consistent. This is just the assertion that any partial ordering on a finite set can be extended to a total ordering. This is true and presumably can be proved by a (perhaps rather mucky?) induction on \mathbb{N} . (ii) is easy.

That was easy. If you are allowed only *propositional* compactness you have to be a bit cleverer. For each $a \neq b$ in X you invent a propositional letter p_{ab} and have axioms $p_{ab} \rightarrow p_{bc} \rightarrow p_{ac}$ and p_{ab} XOR p_{ba} and—if you are lucky enough to know $R(a, b)$ —you adopt p_{ab} as an axiom.

Question 8

Heyting Algebras are the tutelary algebras for constructive logic.

Part (i) asks us to check that K (that is, $a \Rightarrow (b \Rightarrow a)$) is a Heyting tautology.

$$a \Rightarrow (b \Rightarrow a) \quad \text{is} \quad \bigvee \{c : a \wedge c \leq (b \Rightarrow a)\};$$

which is

$$\bigvee \{c : a \wedge c \leq \bigvee \{d : b \wedge d \leq a\}\}. \quad (\text{C})$$

But observe that every c belongs to the underlined set, beco's

$$a \leq \bigvee \{d : b \wedge d \leq a\}$$

and this is true beco's

$$b \wedge a \leq a$$

But this means that the big sup— (C) —is the top element of the algebra.

That wasn't tooooo bad ... now we should do the same for S ...

(iv)

Find a Heyting countermodel for

$$((p \rightarrow q) \rightarrow r) \rightarrow (((q \rightarrow p) \rightarrow r) \rightarrow r)$$

Discussion

This is much less scary than it looks. Observe that the equivalence of $A \rightarrow (B \rightarrow C)$ and $A \wedge B \rightarrow C$ is a Heyting tautology. (I have to admit that—coming from where you are—this is far from obvious. However it is an important and elementary fact.) So we can rearrange our formula (substituting $((p \rightarrow q) \rightarrow r)/A$, $((q \rightarrow p) \rightarrow r)/B$ and r/C) to

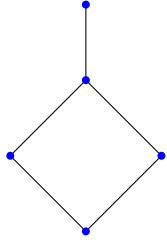
$$(((p \rightarrow q) \rightarrow r) \wedge ((q \rightarrow p) \rightarrow r)) \rightarrow r$$

This is of the form $((A \rightarrow C) \wedge (B \rightarrow C)) \rightarrow C$. Observe that $(A \rightarrow C) \wedge (B \rightarrow C)$ is Heyting-equivalent to $(A \vee B) \rightarrow C$. (Again, altho' this is far from obvious, it is both elementary and important.) So our formula (substituting $p \rightarrow q/A$ and $q \rightarrow p/B$) becomes

$$(((p \rightarrow q) \vee (q \rightarrow p)) \rightarrow r) \rightarrow r$$

It is of the form ' $(A \rightarrow r) \rightarrow r$ ' where ' r ' does not appear in A . What does this say? It says that if r follows from A then r is true. If this allegation is to be Heyting-correct then A had better be a Heyting tautology. For consider: if A is *not* a Heyting tautology then we can cook up an algebra in which it gets some truth-value a other than \top ; since ' r ' does not appear in ' A ' we can assign it a truth-value without reference to A , so give it truth-value a . Under this assignment the two truth-values $[[A]]$ and $[[r]]$ are both a so $[[A \rightarrow r]] = \top$, and $[[(A \rightarrow r) \rightarrow r]] = a \neq \top$, so (on the assumption that A was not a Heyting tautology) ' $(A \rightarrow r) \rightarrow r$ ' was not a Heyting tautology either.

So $(p \rightarrow q) \vee (q \rightarrow p)$ had better be a Heyting tautology, and if we can find a Heyting algebra that refutes it we are in with a chance of finding an algebra that refutes the original formula. As it happens, it isn't, and you can find such an algebra by inspection. It's pretty obvious that $(p \rightarrow q) \vee (q \rightarrow p)$ is going to be good in any Heyting algebra that is a total order and—since it is a classical tautology—it is going to be valid in any Heyting Algebra that is a Boolean algebra. This directs our attention to the smallest Heyting algebra that is neither a total order nor a boolean algebra, namely the five-element Heyting algebra...



The best way to find Heyting countermodels is to use possible world semantics. Have a look at www.dpmms.cam.ac.uk/~tf/chchlectures.pdf. A possible world model will always give rise to a Heyting algebra, and it does so as follows. For each expression ϕ let $[[\phi]]$ be $\{W : W \models \phi\}$. (Think of $[[\phi]]$ as the truth-value of ϕ .) Then the desired Heyting algebra has carrier set $\{[[\phi]] : \phi \in \mathcal{L}\}$, and the partial order is \subseteq . It's a sublattice of the powerset algebra $\mathcal{P}(\mathcal{W})$ where \mathcal{W} is the set of worlds. We write the partial order on \mathcal{W} as $\leq_{\mathcal{W}}$.

Persistence will ensure that $[[\phi]]$ is always $\leq_{\mathcal{W}}$ -upward-closed. This in turn ensures that $[[\phi \wedge \psi]]$ will be $[[\phi]] \cap [[\psi]]$, and that $[[\phi \vee \psi]]$ will be $[[\phi]] \cup [[\psi]]$. We also desire that $[[\phi \rightarrow \psi]]$ will be $[[\phi]] \Rightarrow [[\psi]]$. Now $[[\phi \rightarrow \psi]]$ is $\{W : (\forall W' \geq W)(W' \models A \rightarrow W' \models B)\}$. So certainly $[[\phi \rightarrow \psi]] \cap [[\phi]] \subseteq [[\psi]]$. But we want $[[\phi \rightarrow \psi]]$ to be the \subseteq -largest set of worlds with this property.

inued

While awaiting the continuation we can enjoy Leo Lai's answer...

1. Axiom (K): suppose $(s) = a$ and $(t) = b$. Then $a \wedge b \leq a$ implies $a \leq (ba)$ implies $1 \leq (a(ba))$, so $(s(ts)) = 1$.

Axiom (S): suppose $(r) = a$, $(s) = b$, $(t) = c$, then we need to show $((a(bc))((ab)(ac))) = 1$. By the definition of ,

$$\begin{aligned}
 & 1 \leq ((a(bc))((ab)(ac))) \\
 \text{iff } & (a(bc)) \leq ((ab)(ac)) \\
 \text{iff } & ((ab) \wedge (a(bc))) \leq (ac) \\
 \text{iff } & (a \wedge (ab) \wedge (a(bc))) \leq c
 \end{aligned}$$

To prove the last statement, we have

$$\begin{aligned}
 (a \wedge (ab) \wedge (a(bc))) &= ((a \wedge a) \wedge (ab) \wedge (a(bc))) \\
 &= ((a \wedge (ab)) \wedge (a \wedge (a(bc)))) \\
 &\leq (b \wedge (bc)) \leq c
 \end{aligned}$$

Therefore, (S) is a Heyting tautology.

Given a deduction $(t_1, \dots, t_n = t)$ of t involving only (S) and (K), it will be proven by induction that each t_i is a Heyting tautology. If t_i is an instance of (S) or (K), then the above proves that it is a Heyting tautology. Otherwise, there exists $j, k < i$ such that $t_k = (t_j t_i)$. By the induction hypothesis, $(t_j) = (t_k) = 1$, so $1 = ((t_j) \wedge (t_k)) = ((t_j) \wedge ((t_j)(t_i))) \leq (t_i)$ for all Heyting valuations v . Therefore, t_i is also a Heyting tautology.

2. Let v be a Heyting valuation. Since $0 \leq (q)$, we have $1 \leq (0(q)) = (\perp q)$.
3. Let $T = \{0, a, 1\}$, with $0 \leq a \leq 1$, then T is a Heyting algebra since it is finite and totally ordered. Let $v : \{p, q\} \rightarrow T$, $v(p) = a$, $v(q) = 1$. Then it can be checked that the statement has Heyting valuation a , so it is not a Heyting tautology.
4. Let H be the Heyting algebra of open subsets of $X = (0, 1)$. Let $v : P \rightarrow H$, $v(p) = (0, \frac{1}{2})$, $v(q) = (\frac{1}{2}, 1)$, and $v(r) = X \setminus \{\frac{1}{2}\}$, then

$$\begin{aligned}
(pq) &= (1/2, 1), & (qp) &= (0, 1/2) \\
((pq)r) &= X, & ((qp)r) &= X \\
(((qp)r)r) &= (0, 1/2) \cup (1/2, 1) \\
(((pq)r)((qp)r)r) &= (0, 1/2) \cup (1/2, 1) \neq X
\end{aligned}$$

Therefore, the formula is not a Heyting tautology.

Question 9

Discussion

- (i) The class of finite groups is not axiomatisable. This is because there are arbitrarily large finite groups, so any first-order theory of finite groups would have arbitrarily large finite models and therefore an infinite model (by compactness).
- (ii) The class of infinite groups is axiomatisable; for each $n \in \mathbb{N}$ add an axiom to say there are at least n things. However it's not finitely axiomatisable because, as one can easily show, no finite set of those new axioms implies the whole shebang.
- (iii) Just add an axiom to say there are precisely n things.
- (iv) Suppose there were such a theory; add a constant symbol ' a ' to the language, and infinitely many axioms to say $a^n \neq \mathbb{1}$. By compactness this theory is consistent. So it has a model, which is a group containing an element of infinite order. But this group must be a model of the original theory.

In fact you can do it even without adding a constant symbol to the language: just add, for each n , an axiom to say $(\forall x)(x^n \neq \mathbb{1})$.

- (v) Add to the axioms of group theory the scheme $(\forall x)(x^n \neq \mathbb{1})$.

These should all be comparatively unproblematic, and even a cursory look at old exam questions will bring home to you that this is the sort of thing the examiners expect you to be on top of. If you're not entirely happy about it be sure to work on it until you are.

Actually, as Lev Livnev says, if you want to axiomatise the theory of torsion-free *abelian* groups (instead of (v) as it stands) you can do it finitely, by adding syntax enabling you to say the group is orderable. This is because, by question 7, every torsion free abelian group is orderable.

Question 10

Show that the sentences $(\forall x, y)((x = y) \rightarrow (y = x))$ and $(\forall x, y, z)((x = y) \rightarrow ((y = z) \rightarrow (x = z)))$ are theorems of the predicate calculus with equality. [There is no need to write out formal proofs in full; but you should not expect your supervisor to be satisfied with an argument based on the Completeness Theorem (further exercise: why not?).]

Discussion

It's easy once you remember that in

$$p \rightarrow [x/y]p$$

(which holds when $x = y$) you are not obliged to replace *every* occurrence of ' y ' by ' x '. After all, if $\phi(x, x)$ and $x = y$ then you certainly want to be able to infer $\phi(x, y)$.

There now follows a more detailed answer from Leo Lai:

- | | |
|---|------------------------|
| (1) $(\forall x, y)((x = y)((x = z)(y = z)))$ | (Sub) $\phi = (x = z)$ |
| (2) $((x = y)((x = z)(y = z)))$ | (I) twice |
| (3) $(\forall z)((x = y)((x = z)(y = z)))$ | (Gen) |
| (4) $((x = y)(\forall z)((x = z)(y = z)))$ | (IP) |
| (5) $((\forall z)((x = z)(y = z))((x = x)(y = x)))$ | (I) |
| (6) $((x = y)((x = x)(y = x)))$ | 4 and 5 |
| (7) $(x = x)$ | (Id) and (I) |
| (8) $((x = y)(y = x))$ | 6 and 7 |
| (9) $(\forall x, y)((x = y)(y = x))$ | (Gen) |

where lines 6 and 8 follow by propositional calculus.

Transitivity of equality follows from symmetry and line 1 above.

Question 11

Sorry, chaps, this discussion answer is to the 2015 version of question 11. I promise to update it at some point. This answer is useful too, however, because the two versions are very similar.

- (i) Define the notion of *substructure* of an (Ω, Π) -structure A .
- (ii) Show that if B is a substructure of A and p is a quantifier-free formula of $\mathcal{L}(\Omega, \Pi)$ (with n free variables, say), then $[p]_B = [p]_A \cap B^n$. Give an example to show that this equality may fail if p contains quantifiers.
- (iii) A first-order theory T is called *universal* if its axioms all have the form $(\forall \vec{x})p$ where \vec{x} is a (possibly empty) string of variables and p is quantifier-free. Show that if T is universal, then every substructure of a T -model is a T -model.
- (iv) Similarly, T is called *inductive* if its axioms have the form $(\forall \vec{x})(\exists \vec{y})p$ where p is quantifier-free. Show that if T is inductive, and A is an (Ω, Π) -structure which is the union of a chain of substructures B_i which are T -models, then A is a T -model.
- (v) Which of the theories of question 10 are (axiomatizable as) universal theories? And which are inductive?

[PTJ sez: Naturally-arising counterexamples for (ii) include the centre of a group (the interpretation of $(\forall y)(xy = yx)$) and the group of units of a ring (the interpretation of $(\exists y)(xy = 1)$). Worth emphasizing that in (v) one proves the positive results syntactically (by exhibiting an axiomatization of the required form) and the negative ones model-theoretically (by showing that the properties established in (iii) or (iv) fail for the models of the theory in question); this is a very common phenomenon in model theory. Note also that (vii) is the only non-inductive example in question 10; it was put there in order to provide such an example—in fact very few ‘naturally arising’ first-order theories fail to be inductive.] An obvious exception is set theory!

Discussion

(v) We have to be very careful here. All these theories can be set up as universal theories if we invent enough functions. Even the theory of posets in which every element has a maximal element above it—just add a new function symbol and an axiom $(\forall x)(\forall y)(x \leq f(x) \wedge (y \geq f(x) \rightarrow y = f(x)))$. So what he really means is: as presented here are they universal/inductive etc?

But, hang on—if being a universal theory relies on the class of your models being closed under substructure, how can it come to make a difference to your universal/non-universal status what language you use? The answer is that your choice of language controls which things turn out to be substructures. You can axiomatise group theory in a language with just a single ternary relation “ x times y is z ”, and the axiomatisation you get is certainly not universal. But then a group, as a structure for this language, will have lots of substructures that aren’t groups.

Once you understand why the class of models of an “inductive” (universal-existential) theory is closed under unions of chains, you will be able to see that it’s actually preserved under *directed* unions. (see page 99 for dfn of ‘directed’.)

But here is Leo Lai’s answer to *this* year’s version! (Doctored by me).

1. This can be done by structural induction on ϕ , since the interpretation of predicate symbols, equality, and logical symbols are independent of context.

Let Σ be the empty signature, $A = \{0, 1\}$ considered as a Σ -structure, $B = \{0\}$ a substructure, and $\phi = (\forall y)(x = y)$. Then $\phi_A = \emptyset$, but $\phi_B = \{0\}$.

2. We need to show that if $A \models (\forall \vec{x})\phi$, then $B \models (\forall \vec{x})\phi$ for all substructure B of A , where $\text{FV}(\phi) = \{\vec{x}\}$. By assumption, ϕ does not contain quantifiers, so $\phi_B = \phi_A \cap B^n$, where n is the length of \vec{x} . By the definition of interpretation of universal quantifier, $\phi_A = A^n$, so $\phi_B = B^n$, as required.
3. Again, we consider the case when T consists only of $(\forall \vec{x})(\exists \vec{y})\phi$, where \vec{x} has length n , \vec{y} has length m , and $\text{FV}(\phi) = \{\vec{x}, \vec{y}\}$. Let A be a T -structure, and let $\{A_\alpha | \alpha \in I\}$ be a chain of sub-structures of A . Let $B = \bigcup_{\alpha \in I} A_\alpha$. Let $\vec{b} \in B^n$. The set $\{A_\alpha | \alpha \in I\}$ forms a chain, so there exists an $\alpha \in I$ such that $\vec{b} \subseteq A_\alpha^n$. Then $\vec{b} \in (\exists \vec{y})\phi_{A_\alpha}$ since A_α is a model for T . Let $\vec{y} \in A_\alpha^m \subseteq B^m$ be a witness, then it shows that $\vec{b} \in (\exists \vec{y})\phi_B$. This holds for all $\vec{b} \in B^n$, so $B \models T$.
4. (a) Take the signature $\Sigma = (\Omega, \emptyset)$, where $\Omega = (a, 0, n, m, \mathbb{1})$, $\alpha(a) = \alpha(m) = 2$, $\alpha(0) = \alpha(\mathbb{1}) = 0$, $\alpha(n) = 1$. The class of integral domains has a universal axiomatization with

axioms:

$$\begin{aligned}
&(\forall x, y)(axy = ayx), (\forall x, y, z)(axayz = aaxyz) \\
&(\forall x)(a0x = x), (\forall x)(axnx = 0) \\
&(\forall x, y)(mxy = myx), (\forall x, y, z)(mxmyz = mmyxz) \\
&(\forall x)(m1x = x), (\forall x, y)((mxy = 0)((x = 0) \vee (y = 0))) \\
&(\forall x, y, z)(mxyz = amxymxz), \neg(1 = 0)
\end{aligned}$$

[Mr Lai is using Polish notation to leave out brackets.]

- (b) In the above signature, fields are not universal since the substructure \mathbb{Z} of \mathbb{Q} is not a field. They are inductively axiomatizable by adding the axiom $(\forall x)(\exists y)(\neg(x = 0)(mxy = 1))$. Fields can be made universal by adding a unary operation i to the signature, together with the axiom $(\forall x)(\neg(x = 0)(mxi = 1))$.
- (c) Using the signature for fields, this is not universal, since \mathbb{C} is algebraically closed, but $\mathbb{Q} \subseteq \mathbb{C}$ is not. It is inductive by adding for each n the sentence $(\forall x_1, \dots, x_n)(\exists y)(y^n + x_1y^{n-1} + \dots + x_n = 0)$ to the field axioms.
- (d) Take the signature $\Sigma = (\emptyset, \{\leq\})$, $\alpha(\leq) = 2$ and universal axiomatization

$$\begin{aligned}
&(\forall x)(\leq xx), (\forall x, y, z)(\leq xz(\leq yz \leq xz)) \\
&(\forall x, y)(\neg(x = y)(\leq xy \Leftrightarrow \neg \leq yx))
\end{aligned}$$

- (e) This is not inductive (and hence not universal) in the above signature. Let $A = [\text{set of ordinals below } \omega + 1]$ be given the ordinal ordering. Then each $n \in A \setminus \{\omega\}$ is a substructure of A which is also a model. They form a chain of sub-models in A . However their union is the finite ordinals—which is not a model.

Question 12—Universal Theories⁺

We have a theory T and its offspring the theory T_{\forall} consisting of all the universal sentences that are theorems of T . Let \mathfrak{M} be a model of T_{\forall} , with carrier set M . Add to $\mathcal{L}(T)$ names for every member of M . Add to T all the (quantifier-free) assertions about the new constants that \mathfrak{M} believes to be true. This theory is $T \cup D_M$ (' D ' for 'Diagram'). We want this theory to be consistent. How might it not be? Well, if it isn't, there must be an inconsistency to be deduced from a conjunction ψ of finitely many of the new axioms. This rogue ψ mentions finitely many of the new constants. We have a proof of $\neg\psi$ from T . T knows nothing about these new constants, so clearly we must have a UG proof of $(\forall \vec{x})\neg\psi$. But this would contradict the fact that \mathfrak{M} satisfies every universal consequence of T . So we have established that

THEOREM 3. *For any consistent theory T and any model \mathfrak{M} of T_{\forall} , the theory $T \cup$ the diagram of \mathfrak{M} is consistent.*

Now suppose that T is a theory such that every substructure of a model of T is also a model of T . Let \mathfrak{M} be an arbitrary model of T_{\forall} . We will show that it must be a model of T . We know already from the foregoing that the theory $T \cup D_M$ is consistent, and so it must have a model— \mathfrak{M}^* , say. \mathfrak{M}^* is a model of T , and \mathfrak{M} is a submodel of \mathfrak{M}^* and therefore (by assumption on T) a model of T —as desired.

But all we knew about \mathfrak{M} was that it was a model of the universal consequences of T . So any old \mathfrak{M} that was a model of the universal consequences of T is a model of T . So T is axiomatised by its universal consequences.

So:

THEOREM 4. *If every substructure of a model of T is a model of T then T is axiomatised by its universal consequences.*

This is slightly more useful than you might think. Think about the theory of n -colourable graphs in the language of graphs.. For $n = 2$ finding this theory is an exercise somewhere on one of these sheets, and it can be verified to be a universal theory by inspection. What is the theory of n -colourable graphs for $n \neq 2$? Off the top of my head i don't know (tho' i did find an article in which this is completely explained: <http://www.ams.org/journals/tran/1979-250-00/S0002-9947-1979-0530057-2/S0002-9947-1979-0530057-2.pdf>) but—since every subgraph of an n -colourable graph is n -colourable we do at least know that it is a universal theory.

Question 13

- (i) Show that every countable model of the theory of question 10(vii) of sheet 2 is isomorphic to the ordered set of rational numbers.
- (ii) Is every countable model of first-order Peano arithmetic isomorphic to the set of natural numbers?

Discussion

(i)

I'm assuming that the reader has discovered the back-and-forth construction. I can't be bothered to explain it here, co's it's best done interactively in real time.

It is fairly easy to use the denseness of the rationals to show that every countable linear order can be embedded (in an order-preserving way) into \mathbb{Q} . Think of your countable total order as the members of \mathbb{N} written in a funny order, and then find homes for the natural numbers one by one. That's OK but sadly it isn't quite enough, co's it goes only one way. You might next think "Suppose I have two countable dense linear orders ... I can embed each in the other—so I can then use Cantor-Bernstein!" That doesn't work, beco's Cantor-Bernstein works for *cardinals* not for linear order types—they're far too delicate. (After all, each of the two half-open intervals $(0, 1]$ and $[0, 1)$ embeds in the other but the two are not isomorphic.) So rather than build two embeddings separately, you *interleave* the two constructions in such a way that you construct a single isomorphism—a bijection.

Mind you, there actually *is* a version of Cantor-Bernstein for total orders, even tho' it is no use to us here. If A is iso to a terminal segment of B and B is iso to an initial segment of A then A and B are iso... Actually this is really a theorem about circular orders.

A follow-up thought...

Look at this once you've done sheet 4. Now that you have done ordinals and know what \aleph_1 is—the size of the set of countable ordinals—you might like to think about a generalisation of the fact that by a back-and-forth argument you can show that any two countable dense linear orders without endpoints are isomorphic. There is a theorem that says that any two dense linear orders of

size \aleph_1 without endpoints are isomorphic (by a back-and-forth argument) as long as as they both satisfy a special extra condition. What is that extra condition?

(ii)

QY sez: “No to part (ii). Adjoin to the language a constant c and adjoin to the axioms of Peano arithmetic the sentences $0 < c$, $s(0) < c$, $s(s(0)) < c$, ... to obtain a new theory S . Each finite subset of S has a model, so by compactness S has a model, which is of course infinite. By downward Löwenheim-Skolem, it has a countable model \mathfrak{M} . In \mathfrak{M} there is an element c which is greater than 0 , $S(0)$, $S(S(0))$... but there is no such element in the standard model \mathbb{N} , so \mathfrak{M} is a nonstandard countable model of Peano arithmetic.”

Thanks for this QY, but classroom experience teaches me not leave it at that. Very well, so we have a model of arithmetic with an extra element. But it doesn't stop there. PA proves a whole lot of theorems saying that \mathbb{N} is closed under a lot of operations: $x \mapsto x^2$, $x \mapsto \lceil 22x/7 \rceil$, $x \mapsto \lceil \sqrt{x} \rceil$ and so on. It is probably quite helpful to think of our model as something containing 0 and c and *generated by them*. At its most basic it is a theorem of the arithmetic of \mathbb{N} , after all, that every number has a successor—and that every nonzero number has a predecessor—so we must have $c+1$ and $c-1$. This leads us to the conclusion that c belongs to a copy of \mathbb{Z} stuck on the end of \mathbb{N} . Only one copy...? What about $\lceil 22c/7 \rceil$, $\lceil 355c/133 \rceil$...? In fact a copy of \mathbb{Z} for every rational!

This has the striking (but as far as i know, useless) consequence that all countable nonstandard models of PA are isomorphic as ordered sets. So every countable nonstandard model of PA has order type $\mathbb{N} + \mathbb{Q} \cdot \mathbb{Z}$. You might think that you get *more* than \mathbb{Q} copies of \mathbb{Z} beco's of $\lceil \sqrt{c} \rceil$ but—as noted above, \mathbb{Q} is a maximal countable linear order type so you don't get any further copies of \mathbb{Z} by considering $\lceil \sqrt{c} \rceil$. Of course they aren't all isomorphic as structures for $+$ and \times —beco's arithmetic is incomplete.

I have just learnt the curious fact that every countable nonstandard model of PA is isomorphic to a proper initial segment of itself!

One point one sometimes has to make in this connection is that these wild and woolly things—the nonstandard naturals—living in the desolate marches beyond the standard naturals are absolutely **not** the same wild and woolly things living in the desolate marches beyond ω , namely the countable ordinals. This mistaken identification is a common consequence of over-enthusiastic fault-tolerant pattern matching by beginners.

4.0.3 Sheet 3

Question 1

Is $\{z : \neg(\exists u_1, \dots, u_n)((z \in u_1) \wedge (u_1 \in u_2) \wedge \dots \wedge (u_n \in z))\}$ a set for any n ? [Try to answer this **without** using the axiom of Foundation.]

Discussion

The most useful thing i can do at this juncture is to say the two following things.

- (i) This is an analogue of the Russell paradox, so make sure you understand how the proof of that paradox proceeds.

(ii) You can prove the nonexistence of $\{x : x \notin^n x\}$ (the set of those sets that do not belong to an \in -loop of circumference n) *without any use of set theoretic principles at all!* It is a *theorem of pure logic*. Once you know that you won't be distracted by irrelevancies and will find it easier to locate the correct proof.

Incidentally, for those of you who have a taste for these things, at www.dpmms.cam.ac.uk/~tf/mattgrice.pdf you will find the standard sequent calculus proof of the nonexistence of $\{x : x \notin^2 x\}$. Thanks to Matt Grice. For those of you who care about these things it's worth pointing out that the proof is constructive.

Question 2

Show that the Pair-set axiom is deducible from the axioms of empty set, power set, and replacement.

Two applications of power set to \emptyset gives you $\{\{\emptyset\}, \emptyset\}$ which we then whack with the function class

$$(u = \{\emptyset\} \wedge v = x) \vee (u = \emptyset \wedge v = y)$$

which will give us the pair $\{x, y\}$.

Question 3

Show that if x is a transitive set, then so are $\bigcup x$ and $\mathcal{P}(x)$. Are the converses true?

" x is transitive" is equivalent both to $\bigcup x \subseteq x$ and to $x \subseteq \mathcal{P}(x)$. Observe that both \mathcal{P} and \bigcup are monotone wrt \subseteq , so you get one direction. Crucially \bigcup is *not* injective even tho' \mathcal{P} is, so one of the converses doesn't hold.

Question 4

Question 5

The first thing to note is that it doesn't matter a damn which permutation you use: all the axioms of ZFC (except foundation) are preserved whatever permutation you use, so do not attempt any proof that makes use of particular features of the permutation.

You want to prove $\vdash (\forall \sigma)(\phi \longleftrightarrow \phi^\sigma)$, where ϕ^σ is the result of replacing ' \in ' in ϕ throughout by ' $\in \cdot \sigma$ '.

OK, so you look at ϕ^σ , and you notice that *prima facie* distinct occurrences of a given variable have different prefixes. Variables that never appear to the right of an ' \in ' you say are of level 0, and you don't have a problem with them. Variables that appear to the right of an ' \in ' only when the variable to the left of the \in are of level 1 and you don't have a problem with them—unless they also appear to the left of an \in . Let ' y ' be such a variable. Then we have subformulae like $x \in \sigma(y)$ and $y \in \sigma(z)$ and we have to rewrite the second formula as ' $\sigma(y) \in \sigma(\sigma(z))$ '.

We make the elementary observation that ' $x \in \sigma(y)$ ' is equivalent to ' $\sigma(x) \in \sigma(\sigma(y))$ ' and so can be replaced by it in ϕ where appropriate. σ^*z is $\{\sigma(w) : w \in z\}$ and the function $z \mapsto \sigma^*z$ is of course just yet another permutation. We might find that we have to "lift" σ in this way more than once So the notation ' $j(\sigma)$ ' for this new permutation might come in handy.

The key is to manipulate the formulæ you are dealing with so as to ensure that, for every variable, every occurrence of that variable has the same prefix ...the point being that $(\forall x)(\dots\sigma(x)\dots)$ is equivalent to $(\forall x)(\dots x\dots)$ beco's σ is a permutation.

This is a description of the recursive step in an algorithm for rewriting atomic formulae in such a way that, for each variable, all its occurrences end up with the same prefix, so we can reletter.

The definition of stratifiable for a formula is simply that this algorithm succeeds.

It's now simple to verify that ϕ^σ is equivalent to ϕ as long as ϕ is "stratified". Not all instances of replacement are stratified but it turns out not to matter.

$$(\forall x\exists!y)\phi(x, y) \rightarrow (\forall X)(\exists y)(\forall z)(z \in Y \longleftrightarrow (\exists w)(w \in X) \wedge \phi(w, z))$$

becomes

$$(\forall x\exists!y)\phi^\sigma(x, y) \rightarrow (\forall X)(\exists y)(\forall z)(z \in \sigma(Y) \longleftrightarrow (\exists w)(w \in \sigma(X)) \wedge \phi^\sigma(w, z))$$

We can drop the σ s preceding 'X' and 'Y' to obtain

$$(\forall x\exists!y)\phi^\sigma(x, y) \rightarrow (\forall X)(\exists y)(\forall z)(z \in Y \longleftrightarrow (\exists w)(w \in X) \wedge \phi^\sigma(w, z))$$

which is merely another instance of replacement (as long as σ is a function class). Thus the map (on the syntax) sending each ϕ to ϕ^σ sends every stratified formula ϕ to (something logically equivalent to) ϕ , and sends every instance of replacement to something logically equivalent to another instance.

In any case, all this is explained on pp 201–2 of *Logic, Induction and sets*. (In lemma 105 'NF' should be 'ZF'. This will be corrected in the forthcoming second edition. Other suggestions for corrections welcome).

For Part (iii) go to www.dpmms.cam.ac.uk/~tf/cam_only/FMreadinglist.pdf or perhaps <https://www.dpmms.cam.ac.uk/~tf/basis.pdf>

Questions 6 and 10

6

Define S to be the smallest set a such that $(\emptyset \in a) \wedge (\forall x, y \in a)(x \cup \{y\} \in a)$.

- (i) Explain how the axioms of ZF ensure that such a set exists and is unique.
- (ii) Show that S is closed under \cup .
- (iii) Show that V_ω is closed under \bigcup .
- (iv) Show that S is the smallest set containing \emptyset and closed under taking pairs and unions.

10

Prove that each of the following is an alternative characterization of the set S of question 6:

- (i) V_ω is the class HF of hereditarily finite sets (cf. question 6);
- (ii) V_ω is the class $\{x : \text{TC}(\{x\}) \text{ is finite}\}$ of *strongly hereditarily finite* sets;
- (iii) V_ω is the smallest set containing \emptyset and closed under \mathcal{P} and under formation of arbitrary subsets;
- (iv) $V_\omega = \bigcup\{V_n : n \in \omega\}$ is the ω th stage in the von Neumann hierarchy.

Deduce in particular that the class HF is a set. Is HC a set? If so, does it coincide with V_α for any α ?

The key to these questions is induction, both structural and wellfounded. A good thing to read is Logic, Induction and Sets, ch 2.

Let S be the \subseteq -least set containing \emptyset and closed under $x, y \mapsto x \cup \{y\}$. For every inductively defined set X there is an *in-house* induction principle, which we might as well call “ X ”-induction. S -induction is the principle:

If $F(\emptyset)$ **and**
 $F(y) \rightarrow (\forall z \in S)(F(y \cup \{z\}))$
then $(\forall x \in S)(F(x))$.

How do we know that S even exists? Answer: we use replacement. In fact we *always* use replacement to show that the closure of a set under an operation exists. In this case we consider the function f that sends 0 to the empty set, and thereafter sends n to $\{x \cup \{y\} : x, y \in f(n)\}$. (For pedants out there—and if you are to get on top of this stuff you will eventually need to deal with this very point... How do we write down the function class f ?? I talk more about this in the discussion of question 12 on the next sheet.)

Then we use replacement followed by sumset to obtain $\bigcup f^{\mathbb{N}}$ which (as the reader can verify) is what we want. (The reader might want at this point to think ahead to the later question that asks us how we might use a construction like this to obtain HC , the set of hereditarily countable sets.)

Part (iv) of Question 10 is another inductively defined set, which we will call S' . It, too, has an induction principle. So we can prove by induction that all its members are in S . And we prove by induction on S that all its members are in S' . Hence $S = S'$ by extensionality. Come to think of it there is also the inductively defined set in 6(iv) (which I suppose will have to be S'') and you prove that S'' is identical to the others by a pair of inductions in the same way.

Proving the identity between 10 (i) and 10 (ii) needs induction too, but this time it's \in -induction. Prove by \in -induction that if you are hereditarily finite then you are strongly hereditarily finite, and conversely.

Let's have a look at a couple of these inductions in detail: specifically 6 (ii) and 6(iii).

RTP: S is closed under binary union. (That is, $(\forall x, y \in S)(x \cup y \in S)$.)

Fix $x \in S$ and prove by induction on y that $x \cup y \in S$.

Certainly true for $y = \emptyset$. For the induction suppose $x \cup y \in S$ and deduce $x \cup (y \cup \{z\}) \in S$. $x \cup (y \cup \{z\}) = (x \cup y) \cup \{z\}$. $x \cup y \in S$ by induction hypothesis, and so $(x \cup y) \cup \{z\} \in S$ by closure properties of S .

Next we prove that $x \in S \rightarrow \bigcup x \in S$. Naturally we use S -induction.

Certainly true for $x = \emptyset$. For the induction step assume $\bigcup x \in S$ and infer $\bigcup(x \cup \{y\}) \in S$. Now $\bigcup(x \cup \{y\}) = \bigcup x \cup y$ and we know S is closed under binary union, so $\bigcup(x \cup \{y\}) \in S$.

Several of these sets are not explicitly given as inductively defined. (HF and the set of strongly hereditarily finite sets). For these we can exploit \in -induction. (Note that we have not used \in -induction so far!) We prove by \in -induction that every set that is hereditarily finite is strongly hereditarily finite. It is by \in -induction, too, that we prove that everything in HF is in S or S' or S''

or whatever. This direction really needs \in -induction (aka foundation). If foundation fails then there might be a Quine atom (an $x = \{x\}$ vide Question 5 on this sheet. PTJ calls them autosingletons) and of course there are no Quine atoms in S . We can prove this last fact by $\dots S$ -induction!

To prove that $\text{HF} = V_\omega$ prove by \in -induction on HF that all its members are in V_ω . For the opposite direction prove by induction on n that $V_n \subseteq \text{HF}$.

An easy sequel \dots If AC fails then it might happen that not every infinite set has a countable subset. A set which has no countably infinite subset is *Dedekind-finite* or “D-finite” for short. Show that “the collection of hereditarily D-finite sets” is just another name for the set in question 6. Naturally you are not to use AC in this endeavour!

Question 7

Use the \in -recursion theorem to show that there is a unique function-class $\overline{\text{TC}}$ such that

$$(\forall x)(\overline{\text{TC}}(x) = x \cup \bigcup \{\overline{\text{TC}}(y) : y \in x\}) ,$$

and show that $\overline{\text{TC}}$ coincides with the transitive closure operation as defined in lectures. Why is $\overline{\text{TC}}$ unsatisfactory as a definition of transitive closure?

As Prof Johnstone says, the answer to the last part is, of course, that the existence of transitive closure is used in the proof of the Recursion Theorem.

Question 8

A class M is *transitive* if $(\forall x, y)((x \in y \wedge (y \in M)) \rightarrow (x \in M))$ holds. Show that if M is a transitive class, then the structure (M, \in) satisfies the axiom of extensionality, and that it satisfies each of the empty-set, pair-set and union-set axioms if and only if M is closed under the corresponding finitary operation on V . What more do you need to know about M to get a similar result for the power-set axiom?

[PTJ says: *For the last part, you need ‘super-transitivity’, i.e. the condition that $x \subseteq y \in M$ implies $x \in M$. On the other hand, M can satisfy the power-set axiom without being super-transitive (in which case its ‘power-set operation’ will not coincide with the one on V): the constructible universe (which the students will meet later on) provides an example.*]

Discussion

“finitary operation”??

Most of the axioms of set theory assert that the world of sets is closed under some operation or other. You might think that for a set to be a model of the axiom that says the world of sets is closed under operation blah it is necessary and sufficient for that set to be closed under operation blah. But you’d be wrong! In what follows \mathfrak{M} is the structure consisting of the set M equipped with the membership relation \in .

$\mathfrak{M} \models$ the axiom of pairing iff

$$(\forall x \in \mathfrak{M})(\forall y \in \mathfrak{M})(\exists z \in \mathfrak{M})(\forall w \in \mathfrak{M})(w \in z \longleftrightarrow w \in x \vee w \in y)$$

\mathfrak{M} is closed under the pair set operation iff $(\forall x, y \in \mathfrak{M})(\{x, y\} \in \mathfrak{M})$.

Are these two equivalent? Clearly yes. So far so good.

In contrast $\mathfrak{M} \models$ the axiom of power set iff

$$(\forall x \in \mathfrak{M})(\exists y \in \mathfrak{M})(\forall z \in \mathfrak{M})(z \in y \longleftrightarrow (\forall w \in \mathfrak{M})(w \in z \rightarrow w \in x))$$

Now, since \mathfrak{M} is transitive, the last bit, $(\forall w \in \mathfrak{M})(w \in z \rightarrow w \in x)$, is equivalent to $z \subseteq x$, so the displayed formula simplifies slightly to

$$(\forall x \in \mathfrak{M})(\exists y \in \mathfrak{M})(\forall z \in \mathfrak{M})(z \in y \longleftrightarrow z \subseteq x)$$

\mathfrak{M} is closed under the power set operation iff

$$(\forall x \in \mathfrak{M})(\mathcal{P}(x) \in \mathfrak{M})$$

Are these two equivalent? Clearly not. Reflect that, by Skolemheim, ZF has a countable transitive model \mathfrak{M} . In a countable transitive model every set must be countable. So the thing in \mathfrak{M} that \mathfrak{M} believes to be power set of \mathbb{N} will be a countable set and cannot possibly be the true power set of the naturals.

The way to understand this stuff is to grasp the concept of **restricted quantifier** and Δ_0 -**formula**, and to understand why we restricted our attention to transitive models in the first place. The idea is that if i give you a set x i must also give you all its members—since a set, after all, is nothing more than the set of all its members. So any sensible model with an element x must contain everything in the transitive closure of x as well. Hence our restriction to transitive models only.

We now want to check what conditions we have to put on $\phi(x, y)$ if we are to be confident that the truth value of $\phi(x, y)$ does not depend on the model in which we evaluate it. We say of such ϕ that they are **absolute**. The illustrations PTJ uses are entirely meet for our purpose:

(i) $x = \{y, z\}$ is just

$$y \in x \wedge z \in x \wedge (\forall w \in x)(w = y \vee w = z)$$

and we observe that this can be checked just by looking inside x .

(ii) $x = \bigcup y$ is

$$(\forall w \in x)(\exists z \in y)(w \in y) \wedge (\forall w \in y)(\forall z \in w)(z \in x)$$

(iii) In contrast $x = \mathcal{P}(y)$ is

$$(\forall w \in x)(\forall z \in w)(z \in y) \wedge (\forall w)((\forall u)(u \in w \rightarrow u \in y) \rightarrow y \in x)$$

Observe the difference in the quantifiers. A *restricted* quantifier is one in the style “ $(\forall x \in y) \dots$ ” or “ $(\exists x \in y) \dots$ ”. It turns out that $\phi(x, y)$ is absolute iff all quantifiers within it are restricted. Observe that “ $x = \mathcal{P}(y)$ ” has an unrestricted quantifier in it. We say that a formula is Δ_0 iff all the quantifiers within are restricted. “ $x = \mathcal{P}(y)$ ” is not Δ_0 and is not absolute. If we want power sets to be preserved we have to ensure that the models we deal with not only have all *members* of all of their inhabitants, but also all *subsets* of their inhabitants.

It may be worth noting that the permutation construction of question 5 gives you new sets, but no new subsets of old sets. The power set operation is “absolute” for this construction.

Question 9

If P is a property of sets, a set x is said to be hereditarily P if every member of $\text{TC}(\{x\})$ has property P . Consider the classes HF , HC and HS of hereditarily finite, hereditarily countable and hereditarily small sets (where we call a set small if it can be injected into one of the sets $\omega, \mathcal{P}\omega, \mathcal{P}\mathcal{P}\omega, \dots$): in each case determine which axioms of ZF hold, and which fail, in the structure obtained from the class as in the previous question.

One worry one can banish at the outset: one is assumed to be working in ZFC with foundation, so that foundation is true in all the classes concerned. If $\langle V, \in \rangle$ is wellfounded, so is any substructure of it.

When checking to see whether an axiom ϕ holds in a class H , one key thing to be sure you get right is: *restrict all the variables in ϕ to H !*

Is HC a set?

There is a proof using replacement, where HC_0 is $\{\emptyset\}$ and thereafter HC_α is the set of all countable subsets of $\bigcup_{\beta < \alpha} HC_\beta$. (When does this process close off and why?)

However there is a cuter proof. With the help of countable choice we can find a bijection $\sigma : \{x \subseteq \mathbb{R} : |x| \leq \aleph_0\} \rightarrow \mathbb{R}$.

We can now define a function i by \in -recursion.

$$i(x) =: \sigma(i^{\small{''}}x)$$

We prove by \in -induction that i is defined on all hereditarily countable sets and is injective. This establishes that HC is a set since it is $i^{-1}{}^{\small{''}}\mathbb{R}$, and also establishes that there are at most 2^{\aleph_0} hereditarily countable sets. In fact there are *precisely* 2^{\aleph_0} hereditarily countable sets: for the other direction observe that $V_{\omega+1} \subseteq HC$ and is of size 2^{\aleph_0} .

Observe that $\{V_{\omega+n} : n \in \mathbb{N}\}$ is hereditarily small. It's a countable set of hereditarily small sets. (Prove by induction on n that all the $V_{\omega+n}$ are hereditarily small). But $\bigcup\{V_{\omega+n} : n \in \mathbb{N}\}$ is $V_{\omega+\omega}$ which is not hereditarily small. So HS is not a model of sumset.

Jech proved (JSL 1980) that HC is a set and is of rank ω_2 at most, and he did it without choice. Randall Holmes has shown that—similarly without AC —for *any* X , the class of sets hereditarily smaller than X is always a set. Have a look at www.dpmms.cam.ac.uk/~tf/cam_only/randallhereditary.pdf. It's a nice, idiomatic, delicate piece of set-theoretic combinatorics.

Some detailed thoughts about hereditarily small sets

There is a cluster of independence results concerning the set of hereditarily small sets. There is a certain amount of equivocation going on.

By “hereditarily small” one might mean that

- (i) $TC(x)$ is small;

or one might mean that

(ii) Everything in $TC(x)$ is small;

or even

(iii) x belongs to every set that contains all its small subsets.

(iii) is an inductive definition and supports an induction principle.

By “small” one might mean that $|x| < \beth_\omega$, or one might mean that $(\exists n)(|x| < \beth_n)$. If AC fails, then \beth_ω might not be an aleph and these two assertions about $|x|$ might not be the same. This gives us six sets to consider:

1. $\{x : (\forall y \in TC(\{x\}))(|y| < \beth_\omega)\}$
2. $\{x : (|TC(\{x\})| < \beth_\omega)\}$
3. $\{x : (\forall y \in TC(\{x\}))(\exists n \in \mathbb{N})(|y| < \beth_n)\}$
4. $\{x : (\exists n \in \mathbb{N})(|TC(\{x\})| < \beth_n)\}$
5. $\{x : (\forall y)((\forall z)((z \subseteq y \wedge |z| < \beth_\omega) \rightarrow z \in y) \rightarrow x \in y)\}$
6. $\{x : (\forall y)((\forall z)(\forall n \in \mathbb{N})(z \subseteq y \wedge |z| < \beth_n \rightarrow z \in y) \rightarrow x \in y)\}$

$V_{\omega+\omega}$ is included in all these sets, but is a member of none of them.

I am omitting proofs of the following inclusions, in the belief that they will be obvious to the reader:

- $$\begin{aligned} 4 &\subseteq 2 \subseteq 1; \\ 4 &\subseteq 3 \subseteq 1; \\ 6 &\subseteq 5. \end{aligned}$$

The first class— $\{x : (\forall y \in TC(\{x\}))(|y| < \beth_\omega)\}$ —can be used to prove the independence of sumset from the other axioms of ZF. This is because one of its members is $\{V_{\omega+n} : n \in \mathbb{N}\}$. The same goes for the third set: $\{x : (\forall y \in TC(\{x\}))(\exists n \in \mathbb{N})(|y| < \beth_n)\}$. The fifth and sixth sets, too, contain $\{V_{\omega+n} : n \in \mathbb{N}\}$ but not $V_{\omega+\omega}$ and so prove the independence of sumset.

The second class— $\{x : (|TC(\{x\})| < \beth_\omega)\}$ —on the other hand is clearly a model of sumset: clearly $TC(\bigcup x) \subseteq TC(x)$ so if $|TC(x)| < \beth_\omega$ then $|TC(\bigcup x)| < \beth_\omega$ as well. For the same reason the fourth set $\{x : (\exists n \in \mathbb{N})(|TC(\{x\})| < \beth_n)\}$ will also be a model of sumset.

6 is a model of replacement. If f is a map $6 \rightarrow 6$, then $f^{\ast}x$ is always a set of hereditarily small sets in the appropriate sense. Is it also small itself? $|f^{\ast}x| \leq \beth_n$ so $|f^{\ast}x| < 2^{|f^{\ast}x|} \leq 2^{\beth_n} \leq \beth_{n+1}$ so $f^{\ast}x$ is small too. This doesn't seem to work for 5.

The axiom of power set is a problem. The classes in 3, 4 and 6 will satisfy power set, but for the others we have to assume a certain amount of AC if we want power set to hold. We need $\alpha < \beth_\omega \rightarrow 2^\alpha < \beth_\omega$.

My *Doktorvater* Adrian Mathias claims that 2 is a model of stratified replacement, or even that it is an extension of $V_{\omega+\omega}$ that is elementary for stratified formulæ. How might one prove this, or results like it? In $V_{\omega+\omega}$ one can encode APGs that are pictures of sets in 1-6. If x is a set whose APG is a set in $V_{\omega+\omega}$ then $|TC(\{x\})| \leq \beth_n$ for some $n \in \mathbb{N}$. That is to say, 4 contains precisely the sets whose pictures appear in $V_{\omega+\omega}$. The construction one then performs uses the ideas of

Rieger-Bernays permutations from question 5 of this sheet. Concretise the isomorphism classes of set pictures using Scott's trick. Then reflect that there is a natural "membership" relation on the [isomorphism classes of] set pictures. Swap every [Scott's-trick] isomorphism class with the collection of isomorphism classes that are "members" of it, and extend this to a permutation of the universe by fixing everything else. Look at the Rieger-Bernays permutation model given by this permutation.

Consider the binary relation E on \mathbb{N} defined by: $n E m$ iff the $(n+1)$ st bit (counting from the right) in the binary expansion of m is 1. What can you say about the structure $\langle \mathbb{N}, E \rangle$?

Once you notice that any finite set of naturals can be coded by another natural in this way you quickly realise that $\langle \mathbb{N}, E \rangle \simeq \langle V_\omega, \in \rangle$.

This bijection was first noticed by Wilhelm Ackermann, and is a useful gadget for interpreting arithmetic in set theory and *vice versa*.

this next tho'rt is for after you have learnt about Cantor Normal Forms.

Observe that we can modify Cantor Normal Forms by using base 2 not base ω . (That is to say, just consider the normal function $\alpha \mapsto 2^\alpha$.) Then every ordinal $< \epsilon_0$ has a nice representation as a sum of powers of 2 with exponents less than ϵ_0 . Modify the Ackermann bijection. To what structure does the collection of ordinals below ϵ_0 now correspond?

Question +11

Navigate your way to the **teaching materials** section on my home page, and follow the link to the FM reading group. Then look at the first few pages of the notes you find linked there.

Question 12

Question 13

Let $R \subseteq A \times A$ be a binary relation on a set A . Show that R is well-founded iff there exists a function $h : A \rightarrow \alpha$ for some ordinal, such that $\langle x, y \rangle \in R$ implies $h(x) < h(y)$. Deduce that if every set can be well-ordered, then any well-founded binary relation on a set can be extended to a well-ordering. [Compare question 9(i) on sheet 1.]

(Old version)

Using the axiom of choice, show that every well-founded binary relation on a set can be extended to a well-ordering.

[Hint: given a well-founded relation $r \subseteq a \times a$, first define a 'height function' h_α for some ordinal α , such that $\langle x, y \rangle \in r$ implies $h(x) < h(y)$.]

Discussion

[PTJ's advice to supervisors: h is defined recursively, by a minor variant of the Mostowski collapse; worth remarking that a binary relation is well-founded if and only if it admits a height function to an ordinal. Then choose an arbitrary well-ordering \prec of a , and define a new one by $(x < y) \Leftrightarrow$ either $h(x) < h(y)$ or $(h(x) = h(y) \text{ and } x \prec y)$.]

This is good advice, but there is a cuter way, a *much* cuter way. Use ZL instead of AC. (We know they are equivalent). Given a set A with a wellfounded binary relation R on it, consider the set of those wellorderings S of subsets of A that are *compatible with R* , in the sense that if $R(x, y)$ then $\neg S(y, x)$. This recalls the alternative proof (see the end of the discussion of question 9 on sheet 1, on p. 99). If $R(x, y)$ then y must have higher rank than x and cannot come earlier in any compatible well-ordering. This set of compatible wellorderings can be given a partial order making it into a chain-complete poset, but care is required, since the obvious partial order that we used for q8 on sheet 1 won't do it for us. We say $S \leq S'$ iff S' is an **end-extension** of S . "End-extension"!? You can probably work out what this must mean, so take a minute or so to think about it before reading the next paragraph.

We say S' is an end-extension of S if the graph of S is a subset of the graph of S' **and** all the new things that are being ordered in S' but not in S come *after* all the things in the range of S . The significance of this cunning restriction is that a direct limit of a family of wellorderings under end-extension is another wellordering. That, indeed, was the point of tripos question number 2009-3-16G on page ??, and if you ever really want to properly understand wellorderings and ordinals it is a point worth thinking about (if perhaps not until after the exams). It can even be used to give a recursive characterisation of the class of wellorderings. (Closed under the two operations of (i) plonking an extra thing on the end ["**snoc**"] and (ii) taking (unions of) direct limits under end-extensions).

I'll leave it to you to verify (i) that the poset of R -compatible wellorderings of subsets of A is in fact chain-complete and (ii) that a maximal element must be a compatible wellorder of the whole of A .

However, I think the real value of this question is that it emphasises that a binary relation is wellfounded iff it admits a rank function.

4.0.4 Sheet 4

Question 1

Every limit ordinal is of the form $\omega \cdot \alpha$. Suppose you have a wellordering with no last element. Think about its first point and its limit points. Every such point kicks off a copy of the naturals, a subset of the wellorder of length ω . Thus the wellordering is a union of a concatenation of copies of \mathbb{N} . But that is just to say that its order type is a multiple of ω .

Question 2

[This duplicates Q 2 on Leader sheet 2 2018]

Question 3

“A function-class $F: \mathbf{On} \rightarrow \mathbf{On}$ is called a *normal function* if it is strictly order-preserving (i.e. $\alpha < \beta$ implies $F(\alpha) < F(\beta)$) and continuous at limits (i.e. $F(\lambda) = \sup\{F(\alpha) : \alpha < \lambda\}$ for nonzero limits λ). Prove the following facts about a normal function F :

- (i) $F(\alpha) \geq \alpha$ for all α ;
- (ii) $\sup\{0, F(0), F(F(0)), \dots\}$ is the least ordinal β satisfying $F(\beta) = \beta$;
- (iii) there is a normal function G whose range is exactly the class of ordinals β satisfying $F(\beta) = \beta$.

Find G when F is the function $\beta + (-)$ for some β , and when it is the function $\gamma \cdot (-)$ for some nonzero γ .”

Discussion

(i) You know that F is **strict**: $\alpha < \beta \rightarrow F(\alpha) < F(\beta)$. That means that if $\alpha > F(\alpha)$ ever happened you’d get a descending ω -sequence.

(iii) The temptation here is to define something like: $G(\beta + 1) = \sup\{F^n(\beta) : n \in \mathbb{N}\}$. This makes $G(\alpha)$ the least F -fixed point $> \alpha$. This function is quite useful but it isn’t what we want. Let’s use the letter ‘ H ’ for it instead, so that $H(\beta + 1) = \sup\{F^n(\beta) : n \in \mathbb{N}\}$, taking suprema at limits to make it cts. The correct definition for G is $G(\beta + 1) = H(G(\beta) + 1)$, or—in longhand— $\sup\{F^n(G(\beta) + 1) : n \in \mathbb{N}\}$.

The answers to the last part are

$$\alpha \mapsto \beta \cdot \omega + \alpha.$$

and

$$\alpha \mapsto \gamma^\omega \cdot \alpha.$$

Question 4

Show that every ordinal α has a unique representation (its *Cantor Normal Form*) of the form

$$\alpha = \omega^{\alpha_1} \cdot a_1 + \omega^{\alpha_2} \cdot a_2 + \cdots + \omega^{\alpha_n} \cdot a_n$$

where $n \in \omega$, $\alpha \geq \alpha_1 > \alpha_2 > \cdots > \alpha_n$, and a_1, a_2, \dots, a_n are nonzero natural numbers. [Hint: consider the least β such that $\omega^\beta > \alpha$; can it be a limit?]

You have to exploit a kind of division algorithm for normal functions.
Consider the following diagram:

$$\begin{array}{ccc} \mathcal{P}(On) & \xrightarrow{j(f)} & \mathcal{P}(On) \\ \downarrow \text{sup} & & \downarrow \text{sup} \\ On & \xrightarrow{f} & On \end{array}$$

where ‘ $j(f)$ ’ is the (nonstandard) notation for the function that sends x to $f''x (= \{f(y) : y \in x\})$ that we saw in the discussion of question 5 of sheet 3.

The function f is normal precisely when f is strictly increasing and this diagram commutes. “the value at the limit is the limit of the values”.

There is a discussion of this in www.dpmms.cam.ac.uk/~tf/fundamentalssequence.pdf around p 16.

There is a largest β s.t. $\omega^\beta \cdot n \leq \alpha$. You subtract ω^β from α , and repeat on the result. This can be thought of as an induction on ordinals.

I have just stumbled into a rather nice way of thinking about this stuff. It trades on the idea of absorption-on-the-left: α **absorbs** β **on the left** if $\beta + \alpha = \alpha$. Given a wellordering look for the smallest initial segment that absorbs the corresponding terminal segment (ie, if you added them up the other way round you’d get something smaller). Chop off that initial segment and put it on one side like they tell you to do in all the recipes. Then look at the terminal segment and do the same thing. This decomposes your worder into an ascending finite sequence of initial segments. This decomposition corresponds precisely to CNF....

Question 5

This looks scary but actually it isn’t really. Just do what he tells you.

If you know about Conway numbers you will recognise these operations. If you don’t know about Conway numbers you should read *On Numbers and games*—but not until **after** the exams! (You won’t be able to put it down.)

These operations actually matter. There is a general situation where we have two wellfounded structures and we combine them to make a new wellfounded structure. How does the rank of the new structure depend on the two ranks of the inputs? If the construction is symmetrical in the two inputs then the rank of the new structure will have be obtained by some *symmetric* operation on ordinals.

Leo Lai’s answer (doctored by me)

- (i) Commutativity: Apply well-founded induction on the lexicographical ordering on $ON \times ON$. Given $\alpha, \beta \in ON$, and suppose $\alpha' \oplus \beta' = \beta' \oplus \alpha'$ for all $\alpha', \beta' \in ON$, provided that $((\alpha' < \alpha) \vee (\beta' < \beta))$ holds. Then $\alpha \oplus \beta = \beta \oplus \alpha$ is clear by definition.

Associativity: By definition, $\{\alpha \oplus \beta' | \beta' < \beta\} \cup \{\alpha' \oplus \beta | \alpha' < \alpha\}$ is cofinal in $\alpha \oplus \beta$, and \oplus is increasing in both variables. Therefore,

$$\begin{aligned} (\alpha \oplus \beta) \oplus \gamma &= \sup(\{\delta \oplus \gamma | \delta < \alpha \oplus \beta\} \cup \{(\alpha \oplus \beta) \oplus \gamma' | \gamma' < \gamma\}) \\ &= \sup(\{(\alpha \oplus \beta') \oplus \gamma | \beta' < \beta\} \cup \{(\alpha' \oplus \beta) \oplus \gamma | \alpha' < \alpha\} \\ &\quad \cup \{(\alpha \oplus \beta) \oplus \gamma' | \gamma' < \gamma\}) \end{aligned}$$

A similar formula holds for $\alpha \oplus (\beta \oplus \gamma)$. Apply well-founded induction to the lexicographical ordering on $ON \times ON \times ON$ to get the result as before.

Identity: Do induction on ON . If $0 \oplus \alpha' = \alpha' \oplus 0 = \alpha'$ for all $\alpha' < \alpha$, then

$$\alpha \oplus 0 = \sup(\{(\alpha' \oplus 0)^+ | \alpha' < \alpha\} \cup \{(\alpha \oplus \delta)^+ | \delta < 0\}) = \sup\{(\alpha')^+ | \alpha' < \alpha\}$$

If α is a successor, then it is included in the union, so $\alpha \oplus 0 = \alpha$. Otherwise, if $\delta < \alpha$, then $\delta^+ < \alpha$, so $\alpha \oplus 0 = \sup\{\alpha' | \alpha' < \alpha\} = \alpha$. The result follows by commutativity.

- (ii) If $n = 0$, then $\alpha \oplus 0 = \alpha$ by (i) and $\alpha + 0 = \alpha$ by definition. If $n = 1$, then we prove $\alpha \oplus 1 = \alpha^+ = \alpha + 1$ by induction on α . By definition,

$$\alpha \oplus 1 = \sup(\{(\alpha' \oplus 1)^+ | \alpha' < \alpha\} \cup \{(\alpha \oplus 0)^+\}) = \sup(\{\alpha'^{++} | \alpha' < \alpha\} \cup \{\alpha^+\})$$

by induction hypothesis. If $\alpha' < \alpha$, then $\alpha'^+ \leq \alpha$, so $\alpha'^{++} \leq \alpha^+$. Therefore, the union is at most α^+ . But one of the elements is α^+ , so $\alpha \oplus 1 = \alpha^+$, as required.

Now proceed by induction on n :

$$\alpha \oplus (n+1) = \alpha \oplus (n \oplus 1) = (\alpha \oplus n) \oplus 1 = (\alpha + n) + 1 = \alpha + (n+1)$$

using associativity of \oplus and $+$. This proves the result for all $n < \omega$.

- (iii)-(vi) First note that induction on β shows that $\alpha \oplus \beta \geq \alpha + \beta$ for all $\alpha, \beta \in ON$.

For each ordinal α , let $P(\alpha)$ be the formula

$$(((n < \omega) \wedge (\beta < \omega^\alpha))(\omega^\alpha \cdot n + \beta = \omega^\alpha \cdot n \oplus \beta))$$

and let $Q(\alpha)$ be the formula

$$((n < \omega)(\omega^\alpha \cdot n \oplus \omega^\alpha = \omega^\alpha \cdot (n+1)))$$

We prove by induction on α that $(\forall \alpha)(P(\alpha) \wedge Q(\alpha))$. Note that $P(\alpha') \wedge Q(\alpha')$ for all $\alpha' < \alpha$ implies (vi), restricted to the case when both summands have leading exponent less than α . This in turn implies (iii).

Suppose $P(\alpha')$ and $Q(\alpha')$ holds for all $\alpha' < \alpha$. We first show that $P(\alpha)$ holds. This will be done by well-founded induction on $\langle n, \beta \rangle < \omega \times \omega^\alpha$ with the lexicographic ordering. By definition,

$$\omega^\alpha \cdot n \oplus \beta = \sup\{(\gamma \oplus \beta)^+ | \gamma < \omega^\alpha \cdot n\} \cup \sup\{(\omega^\alpha \cdot n \oplus \beta')^+ | \beta' < \beta\}$$

(The ‘ \cup ’ in the middle of the above formula is a hangover from Leo’s notating these things as von Neumann ordinals. Read it as a ‘max’.)

For the second term, the instance of $P(\alpha)$ with $\langle n, \beta' \rangle$ shows that it is equal to

$$\sup\{\omega^\alpha \cdot n + \beta' + 1 \mid \beta' < \beta\} \leq \sup\{\omega^\alpha \cdot n + \beta \mid \beta' < \beta\} = \omega^\alpha \cdot n + \beta$$

Now consider the first term. If $n = 0$, then it is 0. Otherwise, write γ in Cantor normal form as $\omega^{\alpha_1} \cdot a_1 + \dots$, and call the remaining terms δ . By restricting γ to a cofinal subset of $\omega^\alpha \cdot n$, assume $\alpha_1 = \alpha$, so $a_1 < n$. Instances of $P(\alpha)$ less than $\langle n, \beta \rangle$ show that $\gamma \oplus \beta = (\omega^\alpha \cdot a_1 \oplus \delta) \oplus \beta = \omega^\alpha \cdot a_1 \oplus (\delta \oplus \beta)$. By induction hypothesis, we have $\delta \oplus \beta < \omega^\alpha$, so again by previous instances of $P(\alpha)$, $\gamma \oplus \beta = \omega^\alpha \cdot a_1 + (\delta \oplus \beta) < \omega^\alpha \cdot n \leq \omega^\alpha \cdot n + \beta$. Therefore, $\omega^\alpha \cdot n \oplus \beta \leq \omega^\alpha \cdot n + \beta$. By the earlier observation, this implies $\omega^\alpha \oplus \beta = \omega^\alpha \cdot n + \beta$.

For $Q(\alpha)$, induct on n . If $n = 0$, then both sides equal to ω^α . Otherwise,

$$\omega^\alpha \cdot n \oplus \omega^\alpha = \sup\{(\gamma \oplus \omega^\alpha)^+ \mid \gamma < \omega^\alpha \cdot n\} \cup \sup\{(\omega^\alpha \cdot n \oplus \eta)^+ \mid \eta < \omega^\alpha\}$$

(The ‘ \cup ’ in the middle of the above formula is a hangover from Leo’s notating these things as von Neumann ordinals. Read it as a ‘max’.) The second term equals to $\omega^\alpha \cdot n + \omega^\alpha = \omega^\alpha \cdot (n+1)$ by $P(\alpha)$. For the first term, write $\gamma = \omega^\alpha \cdot a_1 + \delta$ as before. Then by $P(\alpha)$ and induction hypothesis, $\gamma \oplus \omega^\alpha = (\omega^\alpha \cdot a_1 \oplus \delta) \oplus \omega^\alpha = \omega^\alpha \cdot (a_1 + 1) \oplus \delta < \omega^\alpha \cdot (n+1)$. Therefore, $\omega^\alpha \cdot n \oplus \omega^\alpha = \omega^\alpha \cdot (n+1)$.

Given this result, (iii)-(vi) follows trivially.

Question 6

A choice function $g : \mathcal{P}(a) \setminus \{\emptyset\} \rightarrow a$ is said to be *orderly* if

$$g(b \cup c) = g(\{g(b), g(c)\})$$

for all nonempty sets $b, c \subseteq a$. Show that g is orderly iff it is induced by a well-ordering of a . [Hint: if g is so induced, we can recover the ordering by $(x \leq y) \Leftrightarrow (g(\{x, y\}) = x)$.]

Fairly straightforward. To establish wellfoundedness we need to show that any subset a' of a has a \leq -least element. It’s a reasonable bet that the \leq -least element of a should be $g(a)$, and that turns out to be the case.

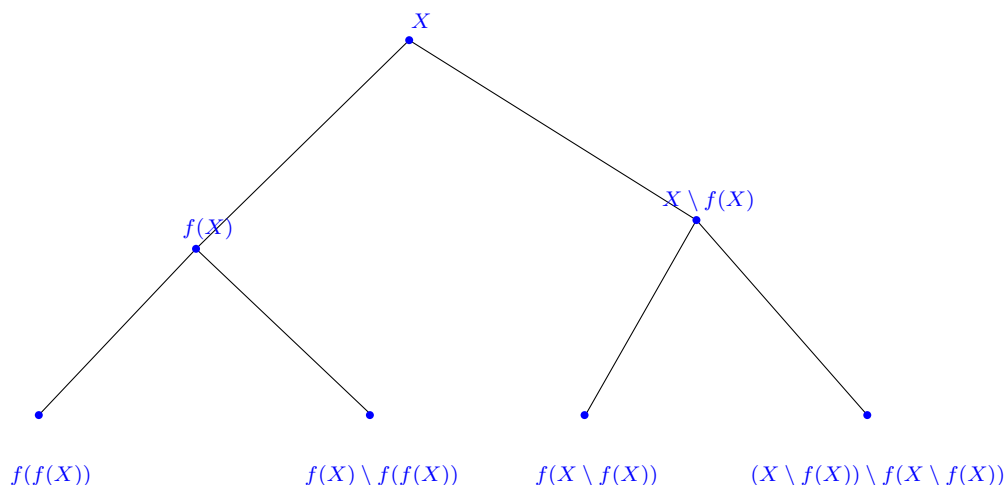
This turns out to have connections with ideas about selection and preference relations from economics. There is a link from my teaching page to some material provided by my former supervisee Peteris Erins, to be found at www.dpmms.cam.ac.uk/tf/cam_only/choice.pdf.

Question –7

One direction was Q9(ii) on sheet 1. For the other direction use Hartogs’ lemma to obtain a wellordered set to inject into your set x . Then, by comparability of cardinals, x injects into this set. But of course anything that is injected into a wellordered set thereby inherits a wellordering from it.

Question 8

Google **Kinna-Wagner**. [You should not use AC for this question, in case you were wondering.]



If you have a function f defined on all subsets Y of X with $|Y| \geq 2$ s.t. $f(Y)$ is always a nonempty proper subset of Y then we can use it to build a gigantic binary tree—the top part of which is illustrated. At the top sits X . Below each point Y you put $f(Y)$ (on the left below Y) and $Y \setminus f(Y)$ (below and to the right). You extend this tree transfinitely downwards by putting below each branch b the set $\bigcap b$, and then pressing on. That way every element of X ends up as the sole inhabitant of the intersection of the end of a branch. (Some branches eventually have empty intersection: we don't worry about those). This gives us a notation for members of X : think about the unique branch to whose intersection x belongs. Think of this branch as a transfinite list of decisions of whether to go left or right at each descending step. Thus a branch is a transfinite (wellordered) sequence of 0s and 1s, which is to say, a subclass of On . But these guys have an obvious lexicographic order, and that order is linear!

Question 9

(i) Show that the statement ‘Every set can be totally ordered’ implies that every family of nonempty finite sets has a choice function.

(ii) ‘Every set has a multiple choice function’ \rightarrow “every totally ordered set can be wellordered.”

Discussion

(i) The trap here is to think that it is sufficient to totally order every finite set in the family. And course every finite set does, indeed, have a total order. (You can prove by ordinary mathematical induction (on \mathbb{N}) that every finite family has a choice function.) But which total order are you going to use? You would have to make an independent choice for each member of the family. That

is (except in the trivial case when your family is finite) *infinitely* many choices, and infinitely-many-choices needs AC. What you have to do is totally order the **union** of all the sets in the family—which you can do beco's of your assumption that every set can be totally ordered. That's *one* choice, which is OK.

(ii) Let X be totally ordered by \leq_X , and suppose there is a multiple choice function f for X . That is to say, if $X' \subseteq X$ is a nonempty subset of X then $f(X')$ is a nonempty finite subset of X' . We will exhibit a choice function on the set of nonempty subsets of X . (That is enough to wellorder X .)

Input Y a subset of X ;
Output the \leq_X -least member of $f(Y)$.

Observe that there is no reason to suppose that the **output** is the \leq_X -least element of Y ; Y might not have a \leq_X -minimal element!

Question 10

Duplicates Leader sheet 2 2018 q 11

Question 11

There is a bijection between \mathbb{R} and $\mathcal{P}(\mathbb{N})$, so every real can be tho'rt of as a set of naturals. There is a bijection between \mathbb{N} and $\mathbb{N} \times \mathbb{N}$ so every natural can be tho'rt of as an ordered pair of naturals. Putting these two facts together means that every real can be thought of as a set of ordered pairs of naturals—which is to say: as a code for a binary relation on \mathbb{N} . Let us say that two reals are *equivalent* if the relations they code are sent to each other by the action of the group of all permutations of \mathbb{N} . Consider the quotient. By considering reals that code equivalence relations on \mathbb{N} we can embed $\mathcal{P}(\mathbb{N})$ into the quotient (two reals encoding equivalence relations are equivalent iff the relations have the same number of equivalence classes of each size—play with the sizes); and we can inject the second number class into the quotient beco's every countable ordinal is the order type [as it were—the equivalence class] of some binary relation on \mathbb{N} .

Question 12

(Tripos II 92206). A subset x of an ordinal α is said to be cofinal if, for every $\beta \in \alpha$, there exists $\gamma \in x$ with $\beta \leq \gamma$. We define the cofinality $\text{cf}(\alpha)$ of α to be the least ordinal β for which there exists an order-preserving map $\beta \rightarrow \alpha$ with cofinal image. Prove that

- (i) for any α , $\text{cf}(\text{cf}(\alpha)) = \text{cf}(\alpha)$;
- (ii) for any limit ordinal α , $\text{cf}(\alpha)$ is an initial ordinal;
- (iii) for any successor ordinal α , $\text{cf}(\omega_\alpha) = \omega_\alpha$;
- (iv) for a nonzero limit ordinal λ , we have either $\text{cf}(\omega_\lambda) < \omega_\lambda$ or $\omega_\lambda = \lambda$.

Show that there is a least ordinal α such that $\omega_\alpha = \alpha$. What is its cofinality?

Discussion

...worth starting off with the observation that it is only part (iii) of this question that needs AC.

I think this question is hard, hard not in the sense of needing clever tricks or new ideas, but hard in the sense of needing to get an awful lot of things straight. A proper discussion will take many pages. I think i might write one up over the summer.

As with all questions about ordinals the only way to really understand what is going on here is to think of ordinals as numbers: lengths of wellorderings. Indeed *all* facts about ordinals can be proved by reasoning about them as isomorphism types of wellorderings³.

One also uses the adjective ‘cofinal’ to describe an unbounded subset of a wellordering. Then the cofinality of an ordinal is the least ordinal that is the order type of an unbounded subset of a wellordering of that length.

On the face of it there are two ways in which one might want to say that one wellordering $\langle B, <_B \rangle$ is at least as long as another wellordering $\langle A, <_A \rangle$. That is to say there are two ways in which one might want to define \leq on ordinals. It could be that

- (i) there is an order-preserving injection from A into B (and this is the idea at play when we consider cofinalities); or it could be that ...
- (ii) there is an orderisomorphism between $\langle A, <_A \rangle$ and an initial segment of $\langle B, <_B \rangle$ (which is how we usually define the order relation on ordinals).

It turns out that these two relations are the same as long as $\langle A, <_A \rangle$ and $\langle B, <_B \rangle$ are wellorderings,⁴ and this makes it much easier to understand cofinalities. For example it is this equivalence that makes it clear that $\text{cf}(\alpha) \leq \alpha$. It also makes it clear that $\text{cf}(\text{cf}(\alpha)) = \text{cf}(\alpha)$, which was part (i).

If $\alpha = \text{cf}(\alpha)$ we say α is **regular**, o/w **singular**. (iii) invites you to prove that all initial ordinals with successor subscripts are regular. You will notice that you used the axiom of choice in proving it. In fact the axiom of choice is necessary: models of ZF have been found in which every transfinite limit ordinal has cofinality ω . Don’t ask.

Part (ii) is nontrivial. I once lectured this at Part III, thinking that it was hard enough for that, but it probably isn’t. I *am* going to lecture it when i lecture this course next year [rather than leave it as an exercise] beco’s it’s hard enough to be worth spelling out. I for one forget how to do it and have to reconstruct it from scratch every time. That’s why i wrote it out! Here is an extract from my next-year’s-notes:

REMARK 4. *Every regular ordinal is initial.*

Proof:

It’s not a particularly deep or important fact but it’s basic and will help you orient yourself. And the proof is idiomatic. Actually we prove the contrapositive.

We need a factoid. Suppose $\langle A, <_A \rangle$ and $\langle B, <_B \rangle$ are (strict) total orders, with $<_A$ a wellorder, and there is a bijection $f : A \rightarrow B$. We are *not* assuming that f is order-preserving! Nevertheless f does have a maximal order-preserving restriction, a rather special one: there is $A' \subseteq A$ s.t $f \upharpoonright A'$ is order-preserving, and $f \upharpoonright A'$ is cofinal (unbounded) in $\langle B, <_B \rangle$.

³Actually that isn’t strictly true, but the falsehood is less misleading than the truth.

⁴This was an exercise on a Hyland example sheet a few years ago, in 2008 i think. A total ordering $\langle A, <_A \rangle$ is a wellordering iff every subordering of it is isomorphic to an initial segment of it. To the best of my knowledge the first published proof is due to your humble correspondent. Who first proved it? God knows. It could even have been Cantor. See page ??.

We obtain A' by recursion on $\langle A, <_A \rangle$. The first member of A' is the bottom element of $\langle A, <_A \rangle$. Thereafter the next member is always the $<_A$ -least element a of A s.t. $f(a) >_B f(a')$ for all $a' <_A a$ that we have already put into A' . Suppose $f''A'$ were bounded in $\langle B, <_B \rangle$. Consider the subset $B' \subseteq B$ consisting of things not dominated by any $f(a)$ for $a \in A'$, and consider the $b \in B'$ s.t. $f^{-1}(b)$ is $<_A$ -minimal. $f^{-1}(b)$ should have been put into A' .

Now suppose β is not an initial ordinal. (As I said, we are proving the contrapositive). Then there is $\alpha < \beta$ s.t. α has as many predecessors as β . Let $\langle A, <_A \rangle$ and $\langle B, <_B \rangle$ (as in the factoid) be the ordinals below α and the ordinals below β respectively. The factoid gives us a set of ordinals cofinal in β whose order type $\leq \alpha < \beta$. So β is not regular. ■

Part (iv) tripped me up, and it may trip you up. Clearly $\text{cf}(\omega_\lambda) \leq \omega_\lambda$, simply beco's $\text{cf}(\alpha) \leq \alpha$ always. Observe—also—that, for limit λ , $\text{cf}(\omega_\lambda) \leq \lambda$; this is beco's the sequence $\langle \omega_\alpha : \alpha < \lambda \rangle$ is a λ -sequence of ordinals whose sup is ω_λ . [In fact for the same reason we have $\text{cf}(\omega_\lambda) \leq \text{cf}(\lambda)$ but we don't need that.]

So, if $\text{cf}(\omega_\lambda)$ is not to be less than ω_λ (i.e. it is precisely ω_λ)—and we know from the above that $\text{cf}(\omega_\lambda)$ is no more than λ —we infer $\omega_\lambda \leq \lambda$, which is to say $\omega_\lambda = \lambda$ as desired. And of course $\text{cf}(\omega_\lambda) = \omega_\lambda$.

One can say more ... the function $\alpha \mapsto \omega_\alpha$ is a normal function and therefore has a fixed point. Look again at your proof (q 3 of this sheet) that all normal functions have fixed points. You do it by a simple-minded iteration ω times. It's worth thinking a little bit about how this construction actually gets carried out, now that we have Set Theory up our sleeves. Let F be a normal function for which we seek a fixed point $\geq \alpha$. Clearly we reach for the supremum of $\{F^n(\alpha) : \alpha \in \mathbb{N}\}$. So we want to know that this last thing is a set. Why might that be? The obvious thing to do is to use replacement, co's the desired set [the set whose sup we are after] is the image of \mathbb{N} in the function that sends n to $F^n(\alpha)$. So: what is the function class that we use? There is a problem because nothing we have done so far authorises us to use natural numbers as *exponents*. You might think we could obtain the graph of this function in the way we obtain \mathbb{N} as an inductively defined set: it would be the intersection of all sets that contain $\langle \alpha, F(\alpha) \rangle$ and are closed under the operation $\langle \beta, x \rangle \mapsto \langle \beta + 1, F(x) \rangle$. Indeed it would; but how do we know that there are any such sets? If we knew that there were such sets then we wouldn't be in the situation of needing replacement, co's we could just obtain the range of the function by a couple of applications of union and separation! No, we need something clever, and I'm not going to keep you in suspense any longer. The function class you need is: " n is related to β iff every set that contains $\langle n, \beta \rangle$ and is closed under $\langle m, F(\gamma) \rangle \mapsto \langle m + 1, \gamma \rangle$ contains $\langle 0, \alpha \rangle$ ".

Sermon over: now we can return to cofinalities.

The fact that you obtain the least fixed point above α by iterating ω times has the effect that the fixed point you construct is the sup of an ω -sequence of approximants, and therefore of course has cofinality ω .

More generally, if f is a normal function we can obtain a fixed point for f of cofinality ω_1 by iterating f through all the countable ordinals. Whenever you hit a fixed point *en route* you just add 1 to it and whack it with f again. In fact you can do this for any ordinal at all: if f is a normal function and α is any (regular) ordinal then f has a fixed point of cofinality α that you obtain by iteration in this way. However it is pretty clear that the fixed points you obtain by these ruses are all singular. The fixed point has cofinality α but it's much bigger than α . Not at all clear how we might prove that every normal function has a *regular* fixed point, and it turns out that this

is a very strong assumption—stronger by far than the consistency of ZF. Consider specifically the normal function before us now, $\alpha \mapsto \omega_\alpha$. Limit initial ordinals ω_λ such that $\text{cf}(\omega_\lambda) = \omega_\lambda$ are said to be *weakly inaccessible*. They are important in Set Theory beco's if λ is such an ordinal then V_λ is a model of *all* the axioms of ZF(C). (You should be able to prove this fact—go back to question 9 on sheet 3.) This means, by the Incompleteness theorem, that we cannot prove the existence of such ordinals in ZF(C).

Question 13

I seem to have mislaid my discussion of this, but here is Leo Lai's!

1. For any two ordinals β and γ , let $(\beta, \gamma) = \{\delta \in \alpha : \beta < \delta < \gamma\}$. Also define $(-\infty, \gamma) = \{\delta \in \alpha : \delta < \gamma\}$. The given basis is $\{(\beta, \gamma) : \beta \in ON \cup \{-\infty\}, \gamma \in ON\}$. This set is closed under finite intersection since $(\beta, \gamma) \cap (\beta', \gamma') = (\beta \cup \beta', \gamma \cap \gamma')$, so it forms a basis for a topology.
2. If $\alpha = 0$, then α is compact. If α is a non-zero limit, then $\alpha = \sup_{\lambda < \alpha} \lambda$. The sets $\{(-\infty, \lambda) : \lambda < \alpha\}$ therefore forms an open cover of α . It does not contain a finite subcover since any finite union of sets of the form is equal to λ for some $\lambda < \alpha$. Therefore, α is not compact.

Now let $\alpha = \beta^+$ be a successor. Let $\{U_i : i \in I\}$ be an open cover of α . Assume inductively that all successor ordinals less than α are compact. There exists $i \in I$ such that $\beta \in U_i$, so there exists a $\gamma < \beta$ such that $(\gamma, \alpha) \subseteq U_i$. The sets $\{U_i : i \in I\}$ restricted to γ^+ also covers γ^+ , so there is a finite subcover. Append U_i to it gives a finite subcover of α . Therefore, α is compact.

3. Let $\{a_n : n < \omega\}$ be a sequence in α . Consider $I = \{n < \omega : (\forall m > n)(a_n \leq a_m)\}$. If it is infinite, then the subsequence of (a_n) indexed by it is a monotonically increasing subsequence. Let $s = \sup_{i \in I} a_i$, then $s < \alpha$ since $\text{cf}(\alpha) > \omega$. Therefore, s is the limit of $\{a_i : i \in I\}$. If I is finite, then for all sufficiently large n , there exists $m > n$ such that $a_m < a_n$. Starting from one such n , we can inductively construct a strictly decreasing sequence of ordinals. Any such sequence must eventually terminate by the axiom of foundation, so I cannot be finite. Therefore, α is sequentially compact.

Let $f : \alpha \rightarrow \mathbb{R}$ be a continuous real function. Suppose f is unbounded; then, for each $n < \omega$, there exists a least $x_n \in \alpha$ such that $|f(x_n)| > n$. By above, the sequence $\{x_n : n < \omega\}$ has a convergent subsequence. However f is unbounded along any subsequence. This contradicts f is continuous. Hence, f is bounded.

4. If $\alpha < \omega_1$, then there exists an order-preserving injection $f : \alpha \rightarrow \mathbb{R}$. Inductively, for each limit ordinal $\lambda < \alpha$, change f to $f - (f(\lambda) - \sup\{f(\lambda') : \lambda' < \lambda\})$ on $\uparrow(\lambda)$, then f preserves both order and limits.

Let d be the pull-back of the standard metric on \mathbb{R} to α . Any open interval (β, γ) for $\beta, \gamma \in \alpha$ maps to the intersection of $(f(\beta), f(\gamma))$ with α , so they are open with respect to d . Conversely, let $U = \{\delta' \in \alpha : d(\delta, \delta') < \epsilon\}$ be a d -open neighborhood around δ . We only need to prove there exists an interval I such that $\delta \in I \subseteq U$, since any other point in U has an open d -ball centered around it contained in U . If $\delta = \eta^+$ is a successor, then $(\eta, \beta^+) = \{\delta\}$, so we are done. Otherwise, δ has cofinality ω by 12(ii), so there exists a sequence $a_n \in \delta$ such that $\sup a_n = \delta$. Since f preserves limits, $\lim f(a_n) = f(\delta)$, so there exists $N < \omega$ such that $|f(\delta) - f(a_N)| < \epsilon$. Then (a_N, β^+) is the required interval.

Suppose $\alpha \geq \omega_1$ is metrizable, then ω_1 is an initial segment of α , so its order topology agrees with the subspace topology induced from $\omega_1 \subseteq \alpha$. Therefore, ω_1 is also metrizable. However, ω_1 is not compact by (i) and sequentially compact by (ii) and 12(iii) (assuming countable choice). This contradicts ω_1 is metrizable as the two concepts are equivalent for metric spaces.