Large Cardinals

Lecturer: Benedikt Löwe

Written by: Paul Minter

Lent Term 2022

These notes are produced entirely from the course I took and my subsequent thoughts. They are not necessarily an accurate representation of the material presented, and may in places have been substantially edited. Please send any corrections to `pdtwm2@cam.ac.uk`; any mistakes are almost certainly my own.

Recommended Books

- Akihiro Kanamori, *The Higher Infinite*.
- Thomas Jech, *Set Theory*.

Contents

1. Introduction

This course is concerned with a certain type of property known as a *large cardinal*. As a wishy-washy definition, by a *large cardinal property* we mean a property of a cardinal κ that implies that

(a) κ is very big;

(b) κ is *so* big that `ZFC` cannot prove the existence of such cardinals.

A *large cardinal* is then a cardinal that has a large cardinal property. A *large cardinal axiom* is an axiom asserting the existence of a cardinal with a certain large cardinal property, i.e., it is an axiom of the form $(\exists\alpha)\Phi(\alpha)$, where Φ is a large cardinal property.

Of course, these are not precise definitions yet. Let us first try to understand what we might mean by a cardinal being "very big" through means of examples.

First recall that if Ord denotes the class of ordinals, then a class-function[1] $F : \mathrm{Ord} \to \mathrm{Ord}$ is called a *normal ordinal operation* if it is both:

(i) monotone increasing, i.e. $\alpha < \beta \Rightarrow F(\alpha) < F(\beta)$;

(ii) continuous in the *order topology*[2], i.e. if λ is a (non-zero) limit ordinal, then $F(\lambda) = \cup_{\alpha<\lambda} F(\alpha)$.

Then we have:

Theorem 1.1. *The class of fixed points for any normal ordinal operation $F : \mathrm{Ord} \to \mathrm{Ord}$ is non-empty and unbounded, i.e. for any ordinal α, there exists an ordinal $\beta \geq \alpha$ for which $F(\beta) = \beta$.*

Proof. First we show that for any non-empty set A of ordinals, we have

$$(\star) \qquad F\left(\bigcup_{\alpha\in A} \alpha\right) = \bigcup_{\alpha\in A} F(\alpha).$$

Of course, as F is increasing we always have for $\alpha \in A$, $F(\alpha) \leq F(\cup_{\beta\in A}\beta)$, and hence $\cup_{\alpha\in A}F(\alpha) \leq F(\cup_{\alpha\in A}\alpha)$. To see the other direction, first suppose that $\cup_{\alpha\in A}\alpha$ is a successor ordinal. Then we would necessarily have $\gamma := \cup_{\alpha\in A}\alpha \in A$, and hence $F(\gamma) \leq \cup_{\alpha\in A}F(\alpha)$, i.e. the other inequality, showing $F(\cup_{\alpha\in A}\alpha) = \cup_{\alpha\in A}F(\alpha)$. If on the other hand $\cup_{\alpha\in A}\alpha$ is a (necessarily non-zero) limit ordinal, then as F is continuous, we would have $F(\cup_{\alpha\in A}\alpha) = \cup_{\beta<\cup_{\alpha\in A}\alpha}F(\beta) = \cup_{\alpha\in A}F(\alpha)$, giving the desired equality and thus proving $(\star)$.

Also, clearly by induction as F is increasing we have $F(\gamma) \geq \gamma$ for all ordinals γ.

Now let us show that for any ordinal α, there is a fixed point β of F with $\beta \geq \alpha$. Indeed, fix such an α, and define a sequence $(\alpha_n)_{n<\omega}$ of ordinals by $\alpha_0 := \alpha$ and $\alpha_{n+1} := F(\alpha_n)$ for all $n \in \omega$. Set $\beta := \cup_{n\in\omega}\alpha_n$; clearly as $\alpha_0 = \alpha$ we have $\beta \geq \alpha$. Moreover, by $(\star)$,

$$F(\beta) = F(\cup_{n\in\omega}\alpha_n) = \cup_{n\in\omega}F(\alpha_n) = \cup_{n\in\omega}\alpha_{n+1} = \beta$$

[1] F cannot be a function of course, as Ord is not a function, but it is a class which behaves like a function.

[2] In the "order topology", successors are isolated points and non-zero limit ordinals are the only possible non-isolated points where a notion of continuity might be needed.

where the last inequality follows from the fact that $\alpha_{n+1} \geq \alpha_n$ for all n, as $F(\gamma) \geq \gamma$ for all ordinals γ. $\square$

Now, if we define inductively on Ord:

- $\aleph_0 := 0$;
- $\aleph_{\alpha+1} :=$ the cardinal successor of $\aleph_\alpha$;
- for λ a non-zero limit ordinal, $\aleph_\lambda := \cup_{\alpha<\lambda}\aleph_\alpha$;

then $F : \mathrm{Ord} \to \mathrm{Ord}$ defined by $F(\alpha) := \aleph_\alpha$ is a normal ordinal operator. Thus, by Theorem 1.1, there are arbitrarily large fixed points of F, i.e. there are arbitrarily large cardinals κ for which

$$\kappa = \aleph_\kappa.$$

These are known as *aleph fixed points*; they must be very large, as since κ is a cardinal, κ is a limit ordinal, and so $\aleph_\kappa = \kappa$ is a limit cardinal. But we know $\aleph_\omega \neq \omega$ (because $\omega \neq 0$); $\aleph_{\omega_1} \neq \aleph_1$ (because $\omega_1 \neq 1$); $\aleph_{\omega_2} \neq \aleph_2$ (as $\omega_2 \neq 2$), and the gap between the indices ω_1 and 1, and ω_2 and 2, is increasing! Similarly, $\aleph_{\aleph_\omega} \neq \aleph_\omega$ (because $\aleph_\omega \neq \omega$).

Therefore, the first aleph fixed point, which is

$$\aleph_{\aleph_{\aleph_{\ddots}}} \equiv \bigcup_{n\in\omega} \aleph_{\aleph_{\ddots_0}}$$

(where on the right-hand side the dots represent n-repetitions), is much bigger than all of these examples above. So it is "very big".

However: it is not "too big" for `ZFC`, i.e. if we define the property AFP of a cardinal κ by:

$$\mathrm{AFP}(\kappa) \;:\Leftrightarrow\; \kappa \text{ is an aleph fixed point}$$

and the axiom AFP by

$$\mathrm{AFP} \;:\Leftrightarrow\; (\exists\kappa)\mathrm{AFP}(\kappa)$$

then we know that `ZFC` $\vdash$ AFP (and this is precisely the Theorem 1.1). So AFP is not a large cardinal property in the wishy-washy sense described previously, despite it being "very large".

1.1. **Inaccessible Cardinals.** To look at another example, recall the notion of *cofinality*:

Definition 1.1. Let λ be a (non-zero) limit ordinal. Then a subset $C \subset \lambda$ is called *cofinal* or *unbounded* if for all $\alpha < \lambda$, there exists $\gamma \in C$ with $\alpha < \gamma$.

Equivalently, if $\lambda = \bigcup C$.

Note that clearly λ is unbounded in λ (for λ any limit ordinal).

Definition 1.2. For λ a (non-zero) limit ordinal, the *cofinality* of λ is:

$$\mathrm{cf}(\lambda) := \min\{|C| : C \text{ is cofinal in } \lambda\}.$$

Clearly the above observation gives that $\mathrm{cf}(\lambda) \leq |\lambda|$ always; thus, for κ a cardinal, $\mathrm{cf}(\kappa) \leq \kappa$. Obviously $\mathrm{cf}(\lambda)$ is a cardinal (as it is a minimum of cardinals), and so therefore if λ is not a cardinal, then as $|\lambda|$ is a cardinal (namely the cardinality of λ) we have $\mathrm{cf}(\lambda) \leq |\lambda| < \lambda$.

Note: It is a simple fact in `ZFC` (as a countable union of sets is countable) that ω_1 is not a countable union of smaller (i.e. countable) ordinals. Thus, in this terminology, we have $\mathrm{cf}(\omega_1) = \omega_1$.

The notion of cofinality can be used to define an interesting pair of properties:

Definition 1.3. A cardinal κ is called *regular* if $\mathrm{cf}(\kappa) = \kappa$; otherwise, i.e. if $\mathrm{cf}(\kappa) < \kappa$, it is called *singular*.

For example, the cardinal $\aleph_\omega$ is singular: indeed,

$$\aleph_\omega := \bigcup\{\aleph_n : n \in \mathbb{N}\}$$

and thus $\{\aleph_n : n \in \mathbb{N}\}$ is cofinal in $\aleph_\omega$, and as these are countable we therefore have $\mathrm{cf}(\aleph_\omega) \leq \aleph_0$ (and therefore in particular we have $\mathrm{cf}(\aleph_\omega) = \aleph_0$, as finite sets are not cofinal in $\aleph_\omega$).

Indeed, in general if $\aleph_\lambda$ is a limit cardinal (i.e. λ is a limit ordinal), then

$$\aleph_\lambda = \bigcup\{\aleph_\alpha : \alpha < \lambda\}$$

and thus $\{\aleph_\alpha : \alpha < \lambda\}$ is a cofinal subset, and hence $\mathrm{cf}(\aleph_\lambda) \leq \mathrm{cf}(\lambda)$.

We have already noted that $\aleph_1$ is a regular cardinal. In fact, it is easy to generalise this to:

Theorem 1.2 (`ZFC`[3])**.** *Every successor cardinal is regular.*

Proof. Let $\kappa = \aleph_{\alpha+1}$. By definition, if $\xi < \kappa$, then there is a surjection from $\aleph_\alpha$ onto ξ. One can then use `AC` to pick such a surjection for each $\xi < \kappa$, i.e. choose for each $\xi < \kappa$ a surjection $\pi_\xi : \aleph_\alpha \to \xi$.

Now, looking for a contradiction, suppose that $\kappa = \bigcup C$, where C is of size $\leq \aleph_\alpha$. Then fix a surjection $\phi : \aleph_\alpha \to C$. Now, define $\pi : \aleph_\alpha \times \aleph_\alpha \to \aleph_{\alpha+1}$ by $(\gamma, \delta) \mapsto \pi_{\phi(\gamma)}(\delta)$. By our assumption that $\kappa = \bigcup C$, we know that π is a surjection. But then we know that $\aleph_\alpha \times \aleph_\alpha$ is in bijection with $\aleph_\alpha$ (this also uses `AC`, and is a standard fact regarding cardinal multiplication in `ZFC`), and hence this gives a surjection from $\aleph_\alpha$ onto $\aleph_{\alpha+1}$, but clearly this is a contradiction by definition. Hence, the contradiction is established and the proof completed. $\square$

Thus, we have seen that: (i) every successor cardinal is regular (so, no singular successor cardinals exist); (ii) all concrete examples of limit cardinals we could come up with were singular. A natural question is then: must every limit cardinal be singular, or do regular limit cardinals exist? This will in fact be our first notion of a "large" cardinal:

Definition 1.4. A cardinal κ is called *weakly inaccessible* if it is a regular limit cardinal.

[3] We will use this to represent when a theorem is going to (heavily) use `AC`.

Let us first show that if κ is weakly inaccessible, then it must be "very big" in some sense. Indeed, let us show that κ must be an aleph fixed point.

Proposition 1.1. *If κ is a weakly inaccessible cardinal, then it must be an aleph fixed point.*

Proof. Suppose κ is weakly inaccessible; in particular, it is a limit cardinal, and so $\kappa = \aleph_\lambda$, where λ is a limit ordinal. We just need to show that $\lambda = \kappa$. But we saw earlier that:

$$\mathrm{cf}(\kappa) \equiv \mathrm{cf}(\aleph_\lambda) \leq \mathrm{cf}(\lambda) \leq \lambda$$

and so if $\lambda < \kappa$, then we would have $\mathrm{cf}(\kappa) \leq \lambda < \kappa$, and thus κ is singular, a contradiction to the assumption that κ was weakly inaccessible. Thus, we must have $\lambda = \kappa$, and hence κ is an aleph fixed point. $\square$

At the moment, we cannot show that `ZFC` does not prove the existence of weakly inaccessible cardinals, which was our second condition of a large cardinal. We will however be able to show this for a slight strengthening of weakly inaccessible. First, define:

Definition 1.5. A cardinal κ is called a *strong limit* if for each $\lambda < \kappa$, we have $2^\lambda < \kappa$.

Note that a limit cardinal was one for which whenever $\aleph_\alpha < \kappa$, we also had $\aleph_{\alpha+1} < \kappa$; loosely speaking, this says that "κ cannot be reached by taking successors". Being a strong limit is saying much more, that κ cannot be reached by taking power sets (as 2^λ is the cardinality of the power set of λ). Clearly, being a strong limit implies being a limit.

We can then define a stronger notion than weakly inaccessible by requiring not only that it is a limit, by that it is a strong limit:

Definition 1.6. A cardinal κ is called (*strongly*) *inaccessible* if it is a regular strong limit cardinal.

We next aim is to prove that `ZFC` cannot prove the existence of inaccessible cardinals. Let us write, for a cardinal κ, $\mathtt{IC}(\kappa)$ to be the term which says "κ is (strongly) inaccessible", i.e."κ is regular and a strong limit". Let us also write `IC` for the sentence $(\exists\kappa)\mathtt{IC}(\kappa)$. We therefore want to show (if `ZFC` is consistent): $\mathtt{ZFC} \nvdash \mathrm{IC}$.

For this, let us first reminder ourselves of the *von Neumann hierarchy* (also known as the *cumulative hierarchy*). We define inductively:

- $V_0 := \emptyset$;
- $V_{\alpha+1} := \mathcal{P}(V_\alpha)$;
- for λ a non-zero limit ordinal, $V_\lambda := \bigcup_{\alpha<\lambda} V_\alpha$.

These have the following properties:

(i) all V_α are transitive sets;

(ii) the V_α are cumulative, i.e. if $\alpha \leq \beta$, then $V_\alpha \subset V_\beta$;

(iii) $V_\alpha \cap \mathtt{Ord} = \alpha$, i.e. if we just take the ordinals in V_α, we get the elements of α.

The most important theorem regarding the von Neumann hierarchy we know (see *Logic and Set Theory*) is that the axiom of foundation/regularity is equivalent to: $(\forall x)(\exists\alpha)(x \in V_\alpha)$. This means that, in ZFC, we can think of the von Neumann hierarchy as generating all sets in a transfinite operation which is guided by all ordinals, and, as it is cumulative, we can think of these ordinals as the birthdates of the sets in our construction; this is the *Mirimanoff rank*, i.e.

$$\rho(x) := \min\{\alpha : x \in V_{\alpha+1}\}.$$

In ZFC, each set has a Mirimanoff rank, and we can prove statements about all sets by induction over the rank.

This is what we know about the V_α as sets. However, we know the V_α are more than sets: we can think of the V_α's themselves as little structures in which set theory happens. This is very closely related to the Mirimanoff rank, as for example if you want to know whether the pairing axiom is true in a V_α, you only have to realise that the pair $\{x, y\}$ has a rank that is one higher than the ranks of x and y, that means that if the ordinal α is closed under the maximum operation and the successor (i.e. $+1$) operation, then the pairing axiom will be true in V_α.

This idea is something which one sees in a first course in logic: for example, which axioms of ZF are satisfied in V_ω (which coincides with the class HF of hereditarily finite sets, i.e. sets which have that the transitive closure $\mathrm{TC}(\{x\})$ is finite), or $V_{\omega+\omega}$? Let us organise these two questions slightly differently: most of the ZFC axioms are of the form "$(\forall x)(\exists y)$" (i.e. "for all x, there exists a power set of x", or "for all x, there exists a union of x") which means that showing that an axiom is true in such a structure is showing a bound on the Mirimanoff rank of the thing for which we claim that it exists. More precisely, proving such an axiom in $(V_\alpha, \in)$ is the same as proving a bound on the rank ρ for the y whose existence is claimed in terms of the x.

For example, look at V_λ for λ a limit ordinal: one can show that $(V_\lambda, \in)$ always satisfies all the axioms of ZFC with the exception of the axiom of infinity and the axiom of replacement. In the case $\lambda = \omega$, one cannot prove the axiom of infinity, but one can prove the axiom of replacement (although the proof is somewhat special, as it really relies on the finiteness of that case), and in the case $\lambda > \omega$, then the axiom of infinity does hold, but one cannot prove the axiom of replacement[4].

So, to summarise: for any limit ordinal $\lambda > \omega$, one gets all the axioms of ZFC in $(V_\lambda, \in)$, except for the axiom of replacement, i.e. $(V_\lambda, \in) \vDash$ ZC (i.e. Zermelo set theory, with AC; F is for Frankel, obviously, but it was Frankel that observed that the axiom of replacement was missing from the axioms).

What we are going to do is use the von Neumann hierarchy in order to prove that the existence of inaccessible cardinals cannot be proved in ZFC. To do this, we will use what we have just observed: if one of the V_λ, $\lambda > \omega$ a limit, is a model of all of ZFC, then we cannot prove the existence of such λ from ZFC alone.

Indeed, we have:

[4]Indeed, in this latter case replacement must really fail, as otherwise you would have shown $V_{\omega+\omega}$ is a model of ZFC which by Gödel's incompleteness theorem would mean you would have proved the inconsistency of ZFC, which would be bad.

Theorem 1.3 (Hausdorff). *If κ is inaccessible, then $V_\kappa \vDash$ ZFC.*

As we have just seen, the proof of this actually reduces to just showing that: κ inaccessible $\Rightarrow$ the axiom of replacement holds in V_κ, i.e. $V_\kappa \vDash$ axiom of replacement.

Before proving this, let us first harvest the fruit and show:

Corollary 1.1. *If ZFC is consistent, then ZFC $\nvdash$ IC.*

Proof. Suppose it did, i.e. suppose that ZFC $\vdash$ IC. Then, Hausdorff's theorem gives that IC implies that there exists a model M with $M \vDash$ ZFC. By Gödel's completeness theorem, we know that the existence of a model of ZFC is equivalent to the consistency of ZFC, i.e. con(ZFC). So, applying modus ponens to the two implications above, we would get: ZFC $\vdash$ con(ZFC). But by Gödel's second incompleteness theorem, we know that this in fact implies that ZFC is inconsistent (and note that this does not contradict anything in the above proof, as "false" implies everything). □

Remark: There is a proof that ZFC does not prove the existence of inaccessible cardinals which does not use Gödel's second incompleteness theorem – we will mention this in a moment.

Before proving Hausdorff's theorem we will need to collect a few more results about the von Neumann hierarchy to understand it a bit better.

Lemma 1.1. *If IC(κ) is true and $\alpha < \kappa$, then $|V_\alpha| < \kappa$.*

Note that this is clearly not the case for non-inaccessible cardinals; e.g., at $\kappa = \omega_1$, we have $\omega + 2 < \omega_1$ by $|V_{\omega+2}| = 2^{2^{\aleph_0}}$ (as $|V_{\omega+1}| = 2^{\aleph_0}$), which is $\geq \omega_1$. So, the lemma is a property which inaccessible cardinals have.

Proof. We prove the lemma by induction on α for fixed κ. The base case is trivial, as $|V_0| = 0 < \kappa$ is always true; so the lemma is true for $\alpha = 0$.

For the successor stage, suppose that $|V_\alpha| < \kappa$. Then:

$$|V_{\alpha+1}| = |\mathcal{P}(V_\alpha)| = 2^{|V_\alpha|} < \kappa$$

where the last inequality comes from the fact that $|V_\alpha| < \kappa$ and κ is a strong limit cardinal. This proves the successor step.

Now suppose that λ is a (non-zero) limit, and that for all $\alpha < \lambda$ we have $|V_\alpha| < \kappa$. Then, by definition of V_λ,

$$|V_\lambda| \equiv \left| \bigcup_{\alpha<\lambda} V_\alpha \right| \leq \bigcup_{\alpha<\lambda} |V_\alpha|$$

and $\{|V_\alpha| : \alpha < \lambda\}$ is a set of ordinals in κ, each of which has size $< \kappa$, and there are $\leq |\lambda| < \kappa$ such ordinals in this set. Hence, as κ is a regular cardinal, we know that $\bigcup_{\alpha<\lambda} |V_\alpha|$ is bounded above in κ. In particular, it is $< \kappa$, and hence we have $|V_\lambda| < \kappa$. Thus, by transfinite induction, we are done. □

Remark: The two conditions of inaccessible cardinals fit exactly with the two cases of the induction, and so this is a good way of describing inaccessibles, i.e. through cardinalities of von Neumann ranks.

So, now we know that von Neumann ranks below κ are small in comparison to κ; this gives us a corollary, because it means that elements of V_κ must be small in comparison to κ:

Corollary 1.2. *If* $\mathtt{IC}(\kappa)$ *is true and* $x \in V_\kappa$*, then* $|x| < \kappa$.

Proof. If $x \in V_\kappa$, then, as κ is a limit ordinal, this means that there is some $\alpha < \kappa$ with $x \in V_\alpha$. But V_α is a transitive set, and so $x \subset V_\alpha$. But then $|x| \leq |V_\alpha| < \kappa$, where the second inequality follows from the above lemma. □

So, inaccessible cardinals are those points in the von Neumann hierarchy where everything that sits in the V_κ level is small compared to the height of the level itself. This now allows us to prove Hausdorff's theorem.

Proof of Hausdorff's Theorem. From our previous discussion, we just need to prove that the axiom of replacement holds in V_κ. As replacement is not very nice, let us strengthen it slightly to something nicer, and then prove the stronger thing. Indeed, strengthen replacement to the following principle, which we denote $(\star)$ and refer to as *second order replacement* (SOR):

"for all functions $F : V_\kappa \to V_\kappa$ and all $x \in V_\kappa$, the image of x under F is also in V_κ, i.e. $F[x] := \{F(y) : y \in x\} \in V_\kappa$."[5]

In particular, the axiom of replacement is the restriction of $(\star)$ to F that are definable, over κ, by a first-order formula, and thus SOR does imply replacement in V_κ. So, if we are able to prove that V_κ satisfies SOR[6] then we will have $V_\kappa \vDash$ axiom of replacement (more discussion on this later).

Let us therefore prove SOR. Fix $F : V_\kappa \to V_\kappa$ and fix $x \in V_\kappa$; we need to show $F[x] \in V_\kappa$. Now, $F[x] = \{F(y) : y \in x\}$. As $F : V_\kappa \to V_\kappa$, clearly, by transitivity of V_κ, if $y \in x$, then $y \in V_\kappa$, and so $F(y) \in V_\kappa$; in particular, $F(y)$ has a Mirimanoff rank below κ, i.e. $\rho(F(y)) < \kappa$. Then, look at the following <u>set</u>:

$$C := \{\rho(F(y)) : y \in X\} \subset \kappa.$$

How large is C? It is at most as large as x, as there is at most one ordinal for each element of x. Thus, $|C| \leq |x| < \kappa$, where the second inequality comes from Corollary 1.2. But then, we have that C is a subset of κ of size $< \kappa$, and so by regularity of κ, C is bounded in κ; say it is bounded by some $\alpha < \kappa$ (see Figure 1 below).

This means that $F[x] \subset V_\alpha$, and so by definition of the von Neumann hierarchy, $F[x] \in V_{\alpha+1}$. But of course $V_{\alpha+1} \subset V_\kappa$, and so $F[x] \in V_\kappa$; this therefore shows V_κ satisfies SOR, and so completes the proof. □

---

[5] i.e. if we have something that looks like a function, then its values on a set are also a set. If you think about the discussion of the replacement axiom, this is what you would like the replacement axiom to say. The reason we did not say this in the actual replacement axiom is because such a statement is not expressible in first order logic, and so instead we needed to talk about those formulas that describe these things and create the axiom of replacement scheme, i.e., the replacement scheme was exactly this, except it only works for those F which are definable by a (first-order) formula.

[6] Note that we cannot write $V_\kappa \vDash$ SOR, as SOR is not a first-order formula.

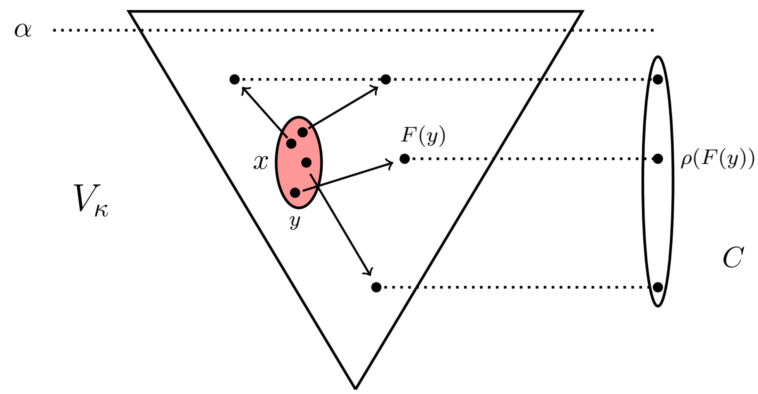

FIGURE 1. Illustration of the proof of Hausdorff's theorem. The set x (in red) contains elements, and we form the set C of the Mirimanoff ranks of their images under F, which we show are upper bounded by some $\alpha < \kappa$. Note that two elements could have distinct images which have the same Mirimanoff ranks (as shown).

This is a pretty big theorem, and the proof was relatively straightforward! We now expand on our previous remark that it is possible to use Hausdorff's theorem to prove that `IC` is not provable in `ZFC` without referring to Gödel's second incompleteness theorem.

Sketch proof of Corollary 1.1 without using Gödel's second incompleteness theorem. What we need to observe is that whether a cardinal is inaccessible is something that is determined essentially by a few levels of the von Neumann hierarchy around the cardinal. One needs to show that, if κ is inaccessible, then for any $\lambda < \kappa$ we have: $\mathtt{IC}(\lambda) \Leftrightarrow V_\kappa \vDash \mathtt{IC}(\lambda)$, i.e. λ is inaccessible if and only if V_κ believes λ is inaccessible – see Example Sheet 1 for this claim.

So, as in our other proof of Corollary 1.1, start by assuming that $\mathtt{ZFC} \vdash \mathtt{IC}$, and so we can choose the *least* inaccessible cardinal, say κ_0. By Hausdorff's theorem we then know that $V_{\kappa_0} \vDash \mathtt{ZFC}$, and hence, as by assumption $\mathtt{ZFC} \vdash \mathtt{IC}$, we have $V_{\kappa_0} \vDash \mathtt{IC}$. Therefore, there is an inaccessible cardinal in V_{κ_0}, i.e. there exists $\lambda < \kappa_0$ such that $V_{\kappa_0} \vDash \mathtt{IC}(\lambda)$. But then by the above claim, if $V_{\kappa_0} \vDash \mathtt{IC}(\lambda)$ then necessarily $\mathtt{IC}(\lambda)$ is true, i.e. we have found an inaccessible cardinal, λ, in the universe which is smaller than the smallest inaccessible cardinal, namely κ_0; this is a contradiction to the minimality of κ_0, which then completes the proof. $\square$

There is a certain slickness in these proofs, where the assumed properties give us exactly what we want; so perhaps there is an equivalence here, and inaccessible cardinals are ones where V_κ is a model of `ZFC`. Let us make this a definition and probe this relationship:

Definition 1.7. A cardinal κ is *worldly* if $V_\kappa \vDash \mathtt{ZFC}$.

Hausdorff's theorem then says: every inaccessible cardinal is worldly. So, could it be that these notions are equivalent, i.e. Hausdorff's theorem is in fact an equivalence? The answer is ultimately no, and that inaccessible is a stronger notion than worldly; however, there is an equivalence hidden in Hausdorff's theorem, namely that inaccessible is not equivalent with V_κ being a model of `ZFC`, but instead is equivalent to `ZFC` with the stronger property of second order replacement (SOR).

So, what is the relationship between inaccessible cardinals and worldly cardinals? It turns out that if κ is inaccessible, then there are many worldly cardinals that are not inaccessible; indeed, the next theorem we will prove is:

Theorem 1.4. *If κ is inaccessible, and $\alpha < \kappa$, then there is λ with $\alpha < \lambda < \kappa$ such that λ is worldly.*

Note that this does not say anything about inaccessibility, but by the same trick we just saw, if you take κ to be the *least* inaccessible, then trivially all of these λ guaranteed by Theorem 1.4 cannot be inaccessible, and thus you find non-inaccessible worldly cardinals.

Note: The role of α in Theorem 1.4 is so that you can find many worldly cardinals, as once you find one such λ, you can just apply the theorem again to λ to find another. So in fact you can find κ-many worldly cardinals below an inaccessible cardinal κ (as κ is necessarily regular).

Before progressing, we need to discuss something somewhat unrelated as it is important to understand a little bit about how worldly cardinals work. If you look at the definition of a worldly cardinal, you see that there is not actually anything a priori which forces κ to be a cardinal; we could say an ordinal α is worldly if $V_\alpha \vDash \mathtt{ZFC}$. It is not obvious that an ordinal being worldly forces the ordinal to be a cardinal. This is something we would like to see first to give us some feeling for how worldly cardinals work.

So, we will first show that if $V_\kappa \vDash \mathtt{ZFC}$, this forces κ to be a cardinal (and so in particular we don't have to say that κ is a cardinal as part of the definition of worldly).

Proposition 1.2. *Every worldly κ is a cardinal.*

Proof. Suppose that κ is worldly, i.e. $V_\kappa \vDash \mathtt{ZFC}$. If κ is not a cardinal, then there exists $\lambda < \kappa$ and a bijection π between λ and κ. Because κ is certainly a limit ordinal[7], we also know that all of $\lambda+1, \lambda+2, \lambda+3$, etc, are $< \kappa$.

But now use the bijection $\pi : \lambda \to \kappa$ to construct a well-order on λ of order type κ in V_κ, namely, set:

$$R := \{(\alpha, \beta) : \pi(\alpha) < \pi(\beta)\}$$

(where, of course, a relation is a set of ordered pairs). By construction, π is an isomorphism between (λ, R) and $(\kappa, \in)$. This means that R is a well-order on λ of order-type κ. But $R \in V_{\lambda+1}$ (as $R \subset V_\lambda$), but this also means that $(\lambda, R) \in V_{\lambda+3}$ (as ordered pairs require two extra levels in the hierarchy), and hence as $V_{\lambda+3} \subset V_\kappa$, we have $(\lambda, R) \in V_\kappa$.

Thus, (λ, R) is a well-order of type κ in V_κ. But V_κ is a model of $\mathtt{ZFC}$ and so in particular V_κ is a model of the representation theorem of well-orders (which says that every well-order is isomorphic to a *unique* ordinal). But (λ, R) can only be isomorphic to κ, which would mean that $\kappa \in V_\kappa$, which is a contradiction; this completes the proof. $\square$

Remark: A minor improvement of the above argument shows that every worldly α is a limit cardinal (see Example Sheet 1).

[7]If not, then $V_{\kappa+1}$ will satisfy "there is a largest ordinal" (namely κ), i.e. $V_{\kappa+1} \vDash$ "there exists a largest ordinal", which is clearly not true in models of $\mathtt{ZFC}$, which it would be by assumption on κ being worldly.

We will now show Theorem 1.4, i.e. that worldly cardinals are not necessarily inaccessible. To do this, we will need some very basic model theorem, and so let us take a small detour.

1.2. **Some Basic Model Theory and Worldly Cardinals.** Let us first recall some basic definitions; L will denote a language throughout.

> **Definition 1.8.** Let M, N be L-structures. We write $M \equiv N$ and say M is *elementary equivalent* to N if and only if for all L-sentences σ, we have $M \vDash \sigma \Leftrightarrow N \vDash \sigma$ (i.e. M, N have the same theories in terms of sentences).

> **Definition 1.9.** Let M, N be L-structures. We write $M \preceq N$ and say M is an *elementary substructure of* N, if and only if $M \subset N$ and for all L-formulas ϕ with n free variables and all n-tuples $a \in M^n$ we have $M \vDash \phi(a) \Leftrightarrow N \vDash \phi(a)$.

Note that usually when we define a notion of equivalence and "less than or equal" (as above), the definition of equivalence is defined by $\leq$ in both directions. But this is not the case with the above two definitions, due to the requirement that $M \subset N$ in the definition of $M \preceq N$. So, if you ever have $M \preceq N$ and $N \preceq M$, this simply just means that the structures are the same (which is clearly much stronger than being elementary equivalent). So: $M \equiv N$ is *not* equivalent to $M \preceq N$ and $N \preceq M$; in fact, being a substructure is stronger than being elementary equivalent, as $M \preceq N \Rightarrow M \equiv N$[8]

The following result is the most important feature/tool of being elementary substructures:

> **Proposition 1.3** (Tarski–Vaught Test (TVT))**.** *Suppose that M is a substructure of N. Then, M is an elementary substructure of N, i.e. $M \preceq N$, if and only if, for any formula $\phi(v, w)$ and $a \in M$, if there is $b \in N$ such that $N \vDash \phi(b, a)$, then there is $c \in M$ such that $N \vDash \phi(c, a)$.*

Loosely speaking, this says: for any formula, if the existential formula in the larger structure is true (i.e. there is $b \in N$), then you can find a witness in the smaller one (i.e. $c \in M$) such that it is still true in the larger structure with this witness.

So, in order to check something is an elementary substructure, the only thing you need to check is that any existential formula which is true in the larger structure is witnessed by something in the smaller one.

Sketch Proof. (See Marter, *Model Theory: An Introduction*, page 45, for a full proof.)

The proof is just by induction on the complexity of formulas, i.e. start with atomic formulas, then propositional connectives and then quantifiers. The only non-trivial step is for the existential quantifier step, and what you need to check is exactly the condition written in the proposition statement. So: (i) atomic formulas and propositional connectives are in fact preserves for all substructures (not just the elementary ones), and so that part is simple; (ii) the only missing step is going from ϕ to $(\exists x)\phi$, and this step needs precisely the assumption given in the proposition. □

This is all the model theory we need for now.

[8] This is because elementary equivalence only talks about sentences, whilst being an elementary substructure talks about parameters from the smaller model as well.

Remark: Since ZFC is a theory (i.e. it consists of sentences), if $V_\kappa \vDash \texttt{ZFC}$ and $V_\lambda \equiv V_\kappa$ are elementary equivalent, then necessarily $V_\lambda \vDash \texttt{ZFC}$. Similarly, if instead $V_\lambda \preceq V_\kappa$, then also $V_\lambda \vDash \texttt{ZFC}$.

In particular, if κ is worldly and $V_\lambda \preceq V_\kappa$, then λ is also worldly. This idea is exactly what we need to prove Theorem 1.4. In fact, we will prove a slightly stronger statement:

Theorem 1.5. *If κ is an inaccessible cardinal, then $\{\lambda < \kappa : V_\lambda \preceq V_\kappa\}$ is unbounded in κ.*

Proof. We use the Tarski–Vaught Test (TVT), which tells us that, for certain $\lambda < \kappa$, we need to show that we can find witnesses in V_λ for every existential formula, i.e. for any $a \in V_\lambda$ for which $V_\kappa \vDash (\exists x)(\phi(x,a))$, then necessarily $V_\kappa \vDash (\exists x)(\phi(x,a) \wedge x \in V_\lambda)$.

The idea of the proof is, for each such formula, we find a suitable λ such that for that formula, the result holds with V_λ. Then we take a union over all the λ and all the formulas, and such that this is an elementary substructure by TVT. So, we will collect witnesses in an ω-iteration procedure.

So fix $\alpha < \kappa$; we can to find λ with $\alpha < \lambda < \kappa$ and $V_\lambda \preceq V_\kappa$. If we want to perform an ω-iteration, we should start at the smallest ordinal largest than α, namely $\alpha + 1$.

So set $\alpha_0 := \alpha + 1$. Suppose now that that $\alpha_i < \kappa$ is already defined. Then, from Lemma 1.1, we know that $|V_{\alpha_i}| < \kappa$. But this also means that, for each n, n-tuples which are parameters from V_{α_i} will still have cardinality $< \kappa$, i.e. if $V_{\alpha_i}^{<\omega}$ denotes the set of finite sequences from V_{α_i}, then $|V_{\alpha_i}^{<\omega}| < \kappa$. Furthermore, we know that there are only countably many formulas, i.e. the set Fml of formulas in the language of set theory is countable. Combining these two bounds, we have

$$|\mathrm{Fml} \times V_{\alpha_i}^{<\omega}| < \kappa.$$

Now, this set $\mathrm{Fml} \times V_{\alpha_i}^{<\omega}$ is sort of the set of instances of the TVT that we need to deal with, i.e. for each formula and each finite sequence from V_{α_i}, we need to find the right witnesses.

What is the witnessing process? Suppose that we have some sequence $a \in V_{\alpha_i}^m$ and $\phi \in \mathrm{Fml}$ with $m+1$ free variables. Then, we need to consider if $V_\kappa \vDash (\exists x)\phi(x,a)$; this can either be true or false. If it is false, we do not care, as the TVT has no requirements in this case. So, if $V_\kappa \nvDash (\exists x)\phi(x,a)$, then simply set $\omega(\phi,a) = 0$ (i.e. set the level where our witness is to be 0; the exact value does not matter, however).

Otherwise, we have $V_\kappa \vDash (\exists x)\phi(x,a)$; hence there is some $\gamma < \kappa$ (as V_κ is a union, as κ is a limit ordinal) such that there is a $c \in V_\gamma$ such that $V_\kappa \vDash \phi(c,a)$. In this case, set $\omega(\phi,a) := \gamma$ (i.e. the level where the witness is).

Since $|\mathrm{Fml} \times V_{\alpha_i}^{<\omega}| < \kappa$, we can look at the set

$$\{\omega(\phi,a) : \phi \in \mathrm{Fml} \text{ with } m+1 \text{ free variables}, a \in V_{\alpha_i}^m\};$$

this is a set of ordinals of size $< \kappa$ (as there is at most one $\omega(\phi,a)$ for each ϕ, a), and so by the regularity of κ, it is bounded, say by some ordinal $\beta < \kappa$. Now we can define:

$$\alpha_{i+1} := \max\{\alpha_i + 1, \beta\};$$

we include the $+1$ to ensure that the sequence of α_i is strictly increasing. This completes the recursive definition of the α_i. Now we define $\lambda := \bigcup_{i\in\mathbb{N}} \alpha_i$. This is clearly an ordinal, which has cofinality $\aleph_0$ (as it is a countable union of smaller ordinals). In particular, $\lambda \neq \kappa$, as the cofinalities differ.

The main claim is that our construction necessarily has that $V_\lambda \preceq V_\kappa$ is an elementary substructure. For this, we just need to check that V_λ satisfies the requirements of TVT. But this is immediate from our construction: indeed, suppose that $V_\kappa \vDash (\exists x)\phi(x,a)$, where $a \in V_\lambda^m$ and ϕ has $m+1$ free variables. Then a is a finite sequence and $\lambda = \cup_{i\in\mathbb{N}}\alpha_i$ is a limit ordinal, and so we can find N sufficiently large such that $a \in V_{\alpha_N}^m$. But now we are in exactly in the setting of our construction when going from α_N to α_{N+1}; therefore, there is a witness for this formula in some γ, and that γ is used to define α_{N+1}. Therefore, there is a witness $c \in V_{\alpha_{N+1}}$ such that $V_\kappa \vDash \phi(c,a)$, which precisely shows the conditions for TVT. This therefore finishes our proof, as this was done for arbitrary $\alpha < \kappa$, and we can repeat with this new $\lambda < \kappa$. $\square$

Remark: As we saw in the proof, for the worldly $\lambda < \kappa$ we construct, we necessarily have $\mathrm{cf}(\lambda) = \aleph_0$. This means that:

(1) The least such worldly cardinal has cofinality $\aleph_0$;

(2) There are in fact unboundedly many worldly cardinals of cofinality $\aleph_0$ below κ, i.e. the set $W^{\aleph_0} := \{\lambda < \kappa : V_\lambda \preceq V_\kappa \text{ and } \mathrm{cf}(\lambda) = \aleph_0\}$ is unbounded in κ (with of course κ inaccessible).

Note that if a worldly cardinal is regular, then (as worldly cardinals are always limit cardinals) then it is weakly inaccessible (see Example Sheet 1). Therefore, there is no chance to prove that every inaccessible cardinal κ will have a smaller *regular* worldly cardinal, i.e. there is no hope to show that any of these constructed worldly $\lambda < \kappa$ are regular.

Why is this? It is because is the Generalised Continuum Hypothesis (GCH) holds, then of course there is no difference between weakly inaccessible cardinals and strongly inaccessible cardinals[9], and certainly we cannot prove that every inaccessible cardinal has a inaccessible cardinal below it (as this will not be true for the least such inaccessible). Thus, in ZFC +GCH, there are inaccessible cardinals with no regular worldly cardinals below.

There is one further open question left from this discussion which will be our next focus, which is: can we get worldly cardinals with higher cofinality, e.g. $\mathrm{cf}(\lambda) = \aleph_1$, or $\aleph_2$, etc. We will see that for every regular cardinal $\kappa' < \kappa$, there are lots of worldly cardinals $< \kappa$ with cofinality κ'.

Why are we interested in this? Remember that our (loose) two conditions to be a large cardinal where (i) to be very big, and (ii) the existence is not provable in ZFC. So, clearly worldly cardinals satisfy (ii) by definition and Gödel's incompleteness theorem, but whether

[9]Since if GCH holds, then for all $\lambda < \kappa$, $2^\lambda = \lambda^+$ is always true, and therefore the notions of limit and strong limit are equivalent, and thus the notions of weakly inaccessible and strongly inaccessible coincide.

they are "very big" is somewhat debatable[10]. Hence, we want to find worldly cardinals which have larger cofinalities, and hence this question.

To start to address this question, we need some further facts from model theory:

(1) If $M_0 \preceq N$, $M_1, \preceq N$, and $M_0 \subset M_1$, then in fact $M_0 \preceq M_1$ (this follows immediately from the definitions).

In light of Theorem 1.5, this means that if κ is inaccessible, then we have lots of worldly pairs (λ, λ') with $\lambda < \lambda' < \kappa$ with $V_\lambda \preceq V_{\lambda'}$.

(2) (Tarski's Chain Lemma). First recall that if $(L, <)$ is a total order and for each $\ell \in L$, M_ℓ is some structure, we call $(M_\ell : \ell \in L)$ an *elementary chain* if for all $\ell < \ell'$ we have $M_\ell \preceq M_{\ell'}$.

Then, Tarski's Chain Lemma says: if $(M_\ell : \ell \in L)$ is an elementary chain and $M := \bigcup_{\ell \in L} M_\ell$, then for each $\ell \in L$ we have $M_\ell \preceq M$.

(The proof of this is just an adaptation of the proof of the Tarski–Vaught Test – see Example Sheet 1.)

We can now prove the following theorem to answer our question:

Theorem 1.6. *If κ is inaccessible and $\mu < \kappa$ is regular, then there exists $\lambda < \kappa$ such that λ is worldly and* $\mathrm{cf}(\lambda) = \mu$.

We remark that, one cannot in general guarantee that $\mu = \lambda$, as then λ would be a regular worldly cardinal, and so it is weakly inaccessible (and thus see the discussion from before concerning GCH).

Proof. Set $W := \{\lambda \in \kappa : V_\lambda \preceq V_\kappa\}$, which we know has size κ from Theorem 1.5. So, take its increasing enumeration, namely $(\lambda_\alpha)_{\alpha\in\kappa}$, where λ_α is the α^{th} element of W.

Now, if $\alpha, \beta \in \kappa$ and $\alpha < \beta$, then $V_{\lambda_\alpha} \preceq V_\kappa$, $V_{\lambda_\beta} \preceq V_\kappa$, and $V_{\lambda_\alpha} \subset V_{\lambda_\beta}$, and so from model theory fact (1) above, we see $V_{\lambda_\alpha} \preceq V_{\lambda_\beta}$. So, if we take any subset of κ, this subset generates an elementary chain, i.e. if $X \subset \kappa$ is an arbitrary subset, then $\{V_{\lambda_\alpha} : \alpha \in X\}$ forms an elementary chain.

So, fix $\mu < \kappa$ regular, and consider the subset $X = \mu$, i.e. the elementary chain $\{V_{\lambda_\alpha} : \alpha < \mu\}$. Then, $\bigcup_{\alpha<\mu} V_{\lambda_\alpha} = V_\lambda$, where $\lambda = \bigcup_{\alpha<\mu} \lambda_\alpha$, and by the Tarski Chain Lemma, we get that $V_{\lambda_\alpha} \preceq V_\lambda$ for any $\alpha < \mu$. So in particular, as $V_{\lambda_\alpha} \vDash \mathtt{ZFC}$ for all $\alpha < \mu$, we see that $V_\lambda \vDash \mathtt{ZFC}$. But by definition of λ, and as μ was regular, we have $\mathrm{cf}(\lambda) = \mu$, and thus we have constructed the desired λ. $\square$

Enough about worldly cardinals for now. Before we move on to other types of large cardinals, let us say more about weakly inaccessible cardinals.

[10] Of course "very big" does not have a precise definition, but a lot of people would say that something with countably cofinality, i.e. can be reached from below by a countable sequence, does not satisfy any reasonable definition of "very big", and so the above constructed worldly cardinals would not be "very big" cardinals. In Kanamori's book, the illustration of large cardinal notions has inaccessible cardinals at the bottom, so many authors would consider inaccessible cardinals to be the smallest large cardinals, and so worldly cardinals are smaller and thus in a grey area.

1.3. **Weakly Inaccessible Cardinals.** We won't be able to see the full proof of the non-provability in ZFC of the existence of weakly inaccessibles, as this relies on a fundamental theorem regarding the existence of inner models of GCH by Gödel which we will not cover (essentially, his proof of the consistency of GCH with ZFC, i.e. $\text{Cons}(\texttt{ZFC}) \to \text{Cons}(\texttt{ZFC}+\texttt{GCH})$). Assuming this however, we can see the rest of the proof. For this, we will need to talk a bit more about models of set theory.

Suppose M, N are models of set theory (e.g. ZFC) such that $M \subset N$ and they use the same element relations, i.e. we are looking at $(M, \in)$ and $(N, \in)$. As this is a language with a single relation, $(M, \in)$ is a substructure of $(N, \in)$, and so in particular atomic formulas remain true if you move from M to N, and as well as if you move from N to M as long as your parameters lie in M.

In general, just being a substructure does not say much in set theory (as there are very few atomic formulas and very simple formulas, such as $(x = 1)$ where "1" is the ordinal, may not retain their meaning if you move to another model). So we need other properties of the models, for example, transitivity[11]. Can we do better for preservation of truth for transitive models?[12]

We can show that a class of formulas larger than the atomic formulas is preserved; this is the class of formulas which is the *closure* of the atomic formulas by propositional connectives and so-called bounded quantification, i.e.

Definition 1.10. Let Λ be a class of formulas. We say that:

(i) Λ is *closed under propositional connectives* if whenever $\phi, \psi \in \Lambda$, then we also have $\phi \wedge \psi, \phi \vee \psi, \neg\phi \in \Lambda$;

(ii) Λ is *closed under bounded quantification* if whenever $\phi \in \Lambda$, then $(\exists x)[(x \in y) \wedge \phi(x)] \in \Lambda$.

(Here, the term y serves as a bound for the variable x; just saying $(\exists x)\phi \in \Lambda$ could be unbounded.)

The class we are interested in is then:

Definition 1.11. The class Δ_0 is the smallest class of formulas that contains the atomic formulas and is closed under both propositional connectives and bounded quantification.

With this, we can now show that if M is transitive in N, then Δ_0 formulas will be preserved; more precisely:

Theorem 1.7. *Suppose M is transitive in N and ϕ is a Δ_0 formula with n free variables. Assume that*[13]$a \in M^n$. *Then:* $M \vDash \phi(a) \iff N \vDash \phi(a)$.

---

[11]Recall that we say M is *transitive* in N if for all $x, y \in N$, if $x \in M$ and $y \in x$, then necessarily $y \in M$. For example, if $\lambda < \kappa$, then V_λ is transitive in V_κ.

[12]A natural question one might ask is whether there are models M, N of ZFC with $M \subset N$ by M is not transitive in N. This is hard, as constructing models of ZFC is hard, but we will see later in the course that example exist where M is not transitive in N.

[13]The parameters need to be in the smaller variables for $M \vDash \phi(a)$ to make sense.

We remark that this property of preservation is also referred to as *absoluteness* of the formula ϕ, i.e. we say that a formula ϕ is *absolute* between M and N if the above holds (for any $a \in M^n$). Thus, this theorem says that Δ_0 formulas are absolute between transitive models of set theory. Note that we do not need to use a precise meaning of "set theory", such as `ZFC`, for this to be true, as it is true in other models of set theory, such as finite set theory.

Proof. Δ_0 is defined to be the smallest class of formulas, and so it is defined inductively by recursion (i.e. first perform the operations between atomic formulae, and then repeat this inductively and take a union), and so it suffices to prove things by induction on the recursion steps.

Thus, we need to prove: (i) that atomic formulas are absolute; (ii) if formulas are absolute than propositional connectives between these formulas are absolute; and (iii) if a formula is absolute, then bounded quantification of it will be absolute. Let us see each of these steps.

Atomic Formulas: We know that these are preserved by all substructures, and so this is immediate.

Propositional Connectives: This is again immediate from the definitions of $\wedge, \vee, \neg$, and is true for all substructures (and so one does not need transitivity here either).

Bounded Quantification: Assume (inductively) that ϕ is absolute between M and N. We know want to show: $(\exists x)(x \in y \wedge \phi)$ is absolute.

$(\Rightarrow)$: Suppose $a \in M$ and $M \vDash (\exists x)(x \in a \wedge \phi)$. This means that we can find $b \in M$ with $M \vDash (b \in a) \wedge \phi$, i.e. $M \vDash (b \in a)$ and $M \vDash \phi(b)$; now, $(b \in a)$ is an atomic formula, and thus is absolute, and moreover ϕ is absolute by the induction assumption. Hence, we have $N \vDash (b \in a)$ and $N \vDash \phi(b)$, i.e. $N \vDash (b \in a) \wedge \phi(b)$, hence $N \vDash (\exists x)(x \in a \wedge \phi)$.[14]

$(\Leftarrow)$: Suppose that $N \vDash (\exists x)(x \in a \wedge \phi)$, where here $a \in M$. By definition, we can therefore find $b \in N$ such that $N \vDash (b \in a) \wedge \phi$, i.e. $N \vDash (b \in a)$ and $N \vDash \phi$. As ϕ is absolute by induction, we know that the latter formula can be transferred to $M \vDash \phi$. However we do not know if $N \vDash (b \in a)$ can be transferred to M as we do not know if $b \in M$; however, by transitivity of M and the fact $a \in M$, we know that $N \vDash (b \in a)$ implies that $b \in M$, and hence as atomic formulas are absolute, $N \vDash (b \in a)$ implies $M \vDash (b \in a)$. Thus, $M \vDash (b \in a) \wedge \phi$, and so $M \vDash (\exists x)(x \in a \wedge \phi)$.

This completes the inductive steps, and so concludes the proof. $\square$

Note: (Trivial example of non-transitive models where Δ_0-formulas are not preserved.) Take $(N, \in) \vDash$ `ZFC` and let $M := \{0, 2\}$ (where $0, 2$ are the ordinals in N). Note that M is not a model of `ZFC` (later we will see a non-transitive $M \subset N$ such that $(M, \in) \vDash$ `ZFC`). Then, look at the formula $\phi(x)$ which says "there is exactly one element of x", i.e.

$$\phi(x) \equiv (\exists u)(u \in x \wedge u = u) \wedge (\forall v)(v \in x \to (\forall w)(w \in x \to v = w)).$$

Note that ϕ is a Δ_0-formula. But then, $(M, \in) \vDash \phi(2)$ as there is precisely one element in 2 in M, namely 0, as $1 \notin M$. By clearly $(N, \in) \nvDash \phi(2)$ (as 2 has two elements, namely 0 and 1). So, ϕ is not absolute for M and N.

[14]Note that this direction does not use transitivity either, and so everything up to here is still true for all substructures.

So, how does Theorem 1.7 help us? It helps because a lot of formulas in set theory are defined by Δ_0 formulas, or are equivalent to Δ_0 formulas in ZFC.

For example, "Being an ordinal": the official definition of being an ordinal is:

$$x \text{ is an ordinal} \iff x \text{ is a transitive set such that } (x, \in) \text{ is a well-order.}$$

This is nota Δ_0 formula as "is a well-order" has existential quantifiers which are not necessarily bounded. But ZFC proves that this is equivalent to[15]:

$$x \text{ is an ordinal} \iff x \text{ is a transitive set and } (x, \in) \text{ is a total order}$$

and this *is* bounded, as one can write a transitivity of a set as:

$$(\forall u)(\forall v)([v \in x \wedge u \in v] \to v \in x)$$

(and note here that there are bounds on the quantifiers, e.g. this is $\forall u$, but then $u \in v$ is a requirement), and a total order is

$$(\forall u)(\forall v)(\forall w)(u, v, w \in x \to [u \in v \wedge v \in w \to v \in w] \wedge [u \notin u] \wedge [u \in v \vee v \in u \vee u = v])$$

and again the quantifiers are bounded. So, in ZFC , being an ordinal is equivalent to a Δ_0 formula, and thus by Theorem 1.7, if M is transitive in N, and $M \vDash \mathtt{ZFC}$, $N \vDash \mathtt{ZFC}$, then:

$$M \vDash (x \text{ is an ordinal}) \iff N \vDash (x \text{ is an ordinal})$$

is true <u>for all $x \in M$</u>. Note that this does not mean this is true for $x \in N$, and so it does <u>not</u> mean $\mathrm{Ord} \cap M = \mathrm{Ord} \cap N$ (for example, it is not true for V_λ, V_κ, with $\lambda < \kappa$.

So, "being an ordinal" is a Δ_0 formula. It should be noted however that some closely related formulas are *not* Δ_0. For example, "being a cardinal" is <u>not</u> a Δ_0-formula. So, let us look at a slightly larger class of formulas where we can describe things like "is a cardinal".

Definition 1.12. A formula is called Σ_1 if it is of the form: $(\exists x)\psi$, where ψ is a Δ_0-formula.

Similarly, a formula is called Π_1 if it is of the form: $(\forall x)\psi$, where ψ is a Δ_0-formula.

Formulas which are Σ_1 or Π_1 are not absolute between transitive models, but they are "half-way" there, in the following senses:

Definition 1.13. Let $M \subset N$ be models of set theory. Then a formula ϕ (with n free variables) is called:

- *downwards absolute* between M and N if for all $x \in M^n$ we have $(N \vDash \phi(x)) \Rightarrow (M \vDash \phi(x))$;
- *upwards absolute* between M and N if for all $x \in M^n$ we have $(M \vDash \phi(x)) \Rightarrow (N \vDash \phi(x))$.

[15] This is because in ZFC the axiom of foundation says that you don't need to say "well-order", because the $\in$-relation is well-founded.

Clearly formulas are absolute if and only if they are both downwards absolute and upwards absolute.

Proposition 1.4. Σ_1 *formulas are upwards absolute and* Π_1 *formulas are downwards absolute for transitive models of* ZFC[16].

Proof. The follows directly from the fact that Δ_0 formulas are absolute (i.e. Theorem 1.7) and $M \subset N$. For example, for the Σ_1 case, if $(M, \in) \vDash (\exists x)\phi$ for some Δ_0 formula ϕ, then there means that there is $a \in M$ with $(M, \in) \vDash \phi(a)$. But because $M \subset N$, this implies that there exists $a \in N$ such that $(M, \in) \vDash \phi(a)$. But by absoluteness of ϕ, this implies that $(N, \in) \vDash (\exists x)\phi$, as desired. The Π_1 case is similar. □

Let us quickly note that the following properties are described, in `ZFC`, by Π_1-formulas:

(1) "κ is a cardinal" – this would be described as "for all $\lambda < \kappa$ and all functions $\lambda \to \kappa$, the function is not a bijection";

(2) "κ is a regular cardinal" – this would be described as "for all $\lambda < \kappa$ and all strictly increasing functions $\lambda \to \kappa$, the range is not cofinal in κ;

(3) "κ is a limit cardinal";

(4) "κ is a strong limit cardinal".

Consequently, we see that if $M \subset N$ is a transitive model of `ZFC` and $k \in M$, and $N \vDash$ "κ is inaccessible", then $M \vDash$ "κ is inaccessible" (and similarly for "κ is weakly inaccessible").

Let us now return to weakly inaccessible cardinals and see how we can prove that `ZFC` $\nvDash$ `WIC`, where `WIC` is $(\exists\kappa)(\kappa$ is a weakly inaccessible cardinal). For this, let us look briefly at `GCH`.

The *generalised continuum hypothesis*, or `GCH`, claims that for all ordinals α:

$$2^{\aleph_\alpha} = \aleph_{\alpha+1}. \tag{GCH}$$

Note that, as:

$$\kappa \text{ is a limit cardinal} \iff (\forall\alpha)(\aleph_\alpha < \kappa \to \aleph_{\alpha+1} < \kappa)$$

(i.e. κ is closed under successor), and

$$\kappa \text{ is a strong limit cardinal} \iff (\forall\alpha)(\aleph_\alpha < \kappa \to 2^{\aleph_\alpha} < \kappa)$$

we readily see that, under `GCH`, these two notions are the same, i.e.

$$\texttt{GCH} \implies (\kappa \text{ is a limit cardinal}) \leftrightarrow (\kappa \text{ is a strong limit cardinal})$$

and so in particular

$$\texttt{GCH} \implies (\kappa \text{ is a weakly inaccessible cardinal}) \leftrightarrow (\kappa \text{ is an inaccessible cardinal}).$$

[16] Whilst `ZFC` is strictly not necessary for the proof, you need a sufficiently strong base theory for the $\Delta_0, \Sigma_1, \Pi_1$ formulas to be interesting, so it is safer to just assume models of `ZFC`.

To prove that ZFC does not prove the existence of weakly inaccessible cardinals (provided ZFC is consistent, of course), we will need Gödel's theorem concerning consistency of ZFC + GCH. For this, we need to strengthen our notion of transitive models to that of *inner models*:

Definition 1.14. Let $M \subset N$ and suppose M is transitive in N and that both are models are ZFC. Then we say that M is an *inner model* of N if $\mathrm{Ord} \cap M = \mathrm{Ord} \cap N$

i.e. if the ordinals in M are the same as the ordinals in N (or, equivalently based on our previous observations, that every ordinal in N is an ordinal in M).

Definition 1.15. Let $M \subset N$ and suppose M is transitive in N and both are models of ZFC. We say that M is *definable in* N if there is a formula Φ such that $(x \in M) \Leftrightarrow N \vDash \Phi(x)$, i.e. if in N, M is defined by some formula.

Then note the following:

Proposition 1.5. *If M is definable in N, then the relation $M \vDash \phi$ is first-order expressible in N, for any formula ϕ.*

Proof. We use the *method of relativisation*: we translate ϕ into a relativised ϕ^M, where all of the quantifiers in ϕ are specified that they only range over M, such that $N \vDash \phi^M$ if and only if $M \vDash \phi$.

Let Φ be the formula which defines M in N. We will do this by replacing all quantifiers in ϕ by a quantifier of the form "there exists something such that Φ holds and ...". More precisely, we do this by recursion on the formulas as follows:

- if ϕ is atomic, then set $\phi^M := \phi$;
- $(\phi \wedge \psi)^M := \phi^M \wedge \psi^M$;
- $(\neg\phi)^M := \neg(\phi^M)$;
- $((\exists x)\phi)^M := (\exists x)(\Phi(x) \wedge \phi)$.

By an easy induction using the definition of ϕ^M, we see that $M \vDash \phi$ if and only if $N \vDash \phi^M$, and thus the statement "ϕ is true in M" is expressible by a formula in N. □

Let us now state Gödel's theorem:

Theorem 1.8 (Gödel, 1938)**.** *If $(N, \in) \vDash$ ZFC, then there is a definable inner model M of N such that $(M, \in) \vDash$ ZFC + GCH.*

We won't prove this (remarkable) theorem: it would be the topic of an entire lecture course itself. This theorem of course implies that Cons(ZFC) implies Cons(ZFC + GCH). The formula defining the inner model in the proof also does not depend on N, i.e. there is a uniform formula defining this particular "Gödel model" in every model of ZFC. This model M is called *Gödel's constructible universe*, usually denoted L. You can even actually strengthen this theorem to assume only that N is a model of ZF, i.e. you do not need AC in N, and so the theorem in fact also proves Cons(ZF) implies Cons(ZFC).

We will prove as a corollary of Gödel's theorem what we are after, i.e.

Corollary 1.3. *Assuming* ZFC *is consistent, then* ZFC $\nvdash$ WIC.

Proof. Take a model of ZFC, i.e. $(N, \in) \vDash$ ZFC. By Gödel's theorem, we know that there is a definable inner model $M \subset N$ such that $(M, \in) \vDash \mathtt{ZFC} + \mathtt{GCH}$.

Looking for a contradiction, let us assume ZFC $\vdash$ WIC. This would then mean that (as weakly inaccessible is a Π_1 formula which is downwards absolute for transitive models of ZFC) $(M, \in) \vDash \mathtt{ZFC} + \mathtt{GCH} + \mathtt{WIC}$. But as we saw before, under GCH the notions of weakly inaccessible and (strongly) inaccessible coincide, and so we would have $(M, \in) \vDash \mathtt{ZFC} + \mathtt{IC}$, i.e. we have a model of ZFC where there is a (strongly) inaccessible cardinal.

We are not quite done here, because what we have now proved is not that there is a *set* model of this, yet, as this M is an inner model of N, which means it has the same ordinals, and so we have not yet constructed a *set* that is a model of ZFC +IC[17].

But we can now redo the proof of Hausdorff's theorem to get a set with this. So, redo the proof of Hausdorff's theorem to show that if κ is an inaccessible cardinal in M, then the von Neumann *set* in M:

$$V_\kappa^M := \{X \in M : M \vDash (\rho(X) < \kappa)\}$$

(note that this is not necessarily the real V_κ intersected with M, but it is just the things which M thinks have Mirimanoff rank $< \kappa$) is first-order expressible in N (by Proposition 1.5) and thus is a set in N; therefore this can be separated from V_κ. But if you look at the proof of Hausdorff's theorem, you prove that this is a model of ZFC, i.e. it gives:

$$N \vDash (V_\kappa^M \vDash \mathtt{ZFC})$$

(in fact, it proves second order replacement). But then this says that N proves the existence of a model of ZFC, i.e. N proves the consistency of ZFC, or alternatively that ZFC proves the Cons(ZFC); but this is a contradiction to the incompleteness theorem. □

1.4. **Digression: A Non-Transitive Submodel of ZFC.** In this section, we will give an example of a non-transitive submodel of ZFC. The idea is essentially the same argument as in the Downward Löwenheim–Skolem Theorem, except we are now being slightly more precise about what we are doing and that we are also going to get an elementary substructure. This is an improvement of the Downward Löwenheim–Skolem Theorem, to obtain not only a countable model of some theory, but in fact a countable elementary substructure of N.

The technique we will use (namely, using the Tarski–Vaught Test) is precisely the one we used in our proof of the existence of worldly cardinals. The idea is to follow the same proof, but now only include the witnesses.

So, start with $M_0 := \emptyset$. If M_i is definable (and countable, which will be true inductively), then consider all $a \in M_i^{<\omega}$ and formulas ϕ, and consider whether $N \vDash (\exists x)\phi(x, a)$. If this is true, let $\omega(\phi, a)$ be a witness (if needed, one can use AC to choose one). If not, let $\omega(\phi, a) = \emptyset$. Then set:

$$M_{i+1} := M_i \cup \{\omega(\phi, a) : \phi \text{ is a formula and } a \in M_i^{<\omega}\}$$

[17]Why do we want a set model? We aim to derive a contradiction by proving Cons(ZFC) in N. But by Gödel's completeness theorem, to do this it suffices (in fact, it is equivalent to) finding a set model of ZFC.

(this is the same as Theorem 1.5, but there we added everything in the von Neumann rank which contained the witness, and here we only add the witness).

If M_i is countable, then so is $M_i^{<\omega}$ (as this is the set of all *finite* sequences from M_i) and as the set of formulas is countable, at each step of this inductive procedure we are only adding countable many witnesses, and so M_{i+1} is countable. Therefore by induction M_i is countable for each i, and hence so is $M := \bigcup_{i\in\mathbb{N}} M_i$ (as it is a countable union of countable sets; again, we are using a form of $\texttt{AC}$ if necessary).

Then, in exactly the same way as in Theorem 1.5, the Tarski–Vaught Test argument gives that $M \preceq N$, i.e. M is a countable elementary substructure of N. In particular, $M \vDash \texttt{ZFC}$. If $N = V_\kappa$ (say), where κ is inaccessible (or even just worldly) then clearly $M \neq N$ (as M is countable).

The main claim is that M *cannot* be transitive in N. Indeed, consider the formula $\phi(x) =$ "x is the smallest uncountable ordinal". Then clearly $N \vDash (\exists x)(\phi(x))$, and this formula *phi* has precisely one witness in N (namely, $x = \aleph_1$, i.e. the $\aleph_1 \equiv \aleph_1^N$ in N). Hence, our construction of M with this ϕ (as the first step of the induction, say) gives that this one witness must be added to M_1, i.e. $\aleph_1 \in M_1 \subset M$.

But M is countable, and so $\aleph_1 \cap M$ is countable, i.e. only countably many elements of $\aleph_1$ are in M, and so $\aleph_1 \not\subset M$; but this is precisely what it means to be not transitive, to have an element which is not a subset. So, M is not transitive in N. $\square$

> **Note:** $\aleph_1$ is not special in this construction: we also have formulas defining $\aleph_2, \aleph_3, \aleph_\omega$, etc: all these ordinals exist in N and have unique witnesses, and so they all get put into M, but a lot of stuff inbetween cannot get put in M, because they are just too big and M is too small. So this M constructed is incredibly sparse, but it is an elementary substructure of N, and so not only is it a model of $\texttt{ZFC}$, but it is indistinguishable from N.

2. Larger Large Cardinals

We now introduce some definitions of other large cardinals. The ones we will look at (in order of decreasing strength):

- strongly compact (s.c.) cardinals;
- measurable cardinals;
- weakly compact (w.c.) cardinals;
- (inaccessible cardinals).

Our goal for the moment will then be the implications above, namely:

$$\text{strongly compact} \implies \text{measurable} \implies \text{weakly compact} \implies \text{inaccessible}.$$

Of course, these implications by themselves do not tell us that the notions are getting stronger, as they could be completely equivalent: we will get to this after establishing the above implications (and these non-implications are the more interesting parts).

Let us start with measurable cardinals. We first will see how this relates to the "measure problem" from measure theory, first asked by Lebesgue and resolved by Vitali in 1905.

2.1. **Measurable Cardinals.** The *measure problem* (for the unit interval, $[0,1]$) asks the following:

"Is there a function $\mu : \mathcal{P}([0,1]) \to [0,1]$ such that:
- $\mu([0,1]) = 1$;
- $\mu(\emptyset) = 0$;
- μ is translation invariant, i.e. $\mu(A+x) = \mu(A)$, where $+x$ refers to a translation (modulo 1);
- μ is σ-additive, i.e. if $\{A_i : i \in \mathbb{N}\}$ is a family of pairwise disjoint sets, then:

$$\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}} \mu(A_i) \quad ?\text{"}$$

This was answered negatively by Vitali in 1905. More precisely, Vitali showed that `AC` implies that there is always a non-measure set; even more precisely, Vitali showed that if there is a basis for $\mathbb{R}$ as a $\mathbb{Q}$-vector space, then the answer is no.

The measure-theoretic resolution to this problem is to say that the problem lies in trying to measure *every* subset of $[0,1]$, and so as this is not possible (as "bad sets" exists) we replace this condition by defining the measure not on all of $\mathcal{P}([0,1])$, but only on certain subsets of $\mathcal{P}([0,1])$ known as σ-algebras, which are precisely the collection of subsets of $\mathcal{P}([0,1])$ where these "measure" functions exists. So essentially, in measure theory one gives up on $\text{dom}(\mu) = \mathcal{P}([0,1])$, and works with appropriate σ-algebras instead.

This is something we will not look at. Instead, in set theory the question of interest is: under what circumstances can you get a function as originally asked by Lebesgue? For example:

- can you find such a function if you weaken some of the conditions?
- can you prove that no such μ exists without using `AC`?

- is this something special about $[0,1]$? What if you change $[0,1]$ to some other set to get such a μ?

The first generalisation was due to Banach: if one drops the condition of translation invariance[18], then you could have trivial solutions, e.g. for $[0,1]$ you could set

$$\mu(A) := \begin{cases} 1 & \text{if } 0 \in A; \\ 0 & \text{if } 0 \notin A \end{cases}$$

and this is σ-additive, and is a solution to the measure problem without translation invariance. So, in order to exclude these trivial solutions, Banach said that we should at least assume that singletons get measure 0 also. So, Banach said that μ is *non-trivial* if for all $x \in [0,1]$, we have $\mu(\{x\}) = 0$ (and therefore, by σ-additivity, all countable sets get measure 0).

This now allows us to define this question for arbitrary sets, as we have removed the only condition which depended on the geometry of $[0,1]$, which wouldn't hold (or even make sense) for arbitrary sets. However, one further abstraction step can be done (which is due to Ulam):

Definition 2.1. An *Ulam measure* on a set X is a function $\mu : \mathcal{P}(X) \to \{0,1\}$ that is non-trivial (i.e. if $x \in X$, then $\mu(\{x\}) = 0$), σ-additive, and obeys $\mu(X) = 1$, $\mu(\emptyset) = 0$.

Note that because an Ulam measure μ can only take values in $\{0,1\}$, σ-additivity implies that disjoint sets cannot both have value 1, i.e. if $A \cap B = \emptyset$, then at most one of A, B can have measure 1 (where by "measure of A" we mean the value $\mu(A)$).

This can then be used to define a "large cardinal" notion:

Definition 2.2. A cardinal κ is called *Ulam measurable* if there is an Ulam measure on κ.

We won't really talk about Ulam measurable cardinals (see Example Sheet 2) because we want to strengthen this notion once more. The reason for this is that this notion of σ-additivity was only really used before because we were on $\mathbb{R}$ (and countable union are essentially the fundamental defining property of $\mathbb{R}$) but in this more general setting, σ-additivity is no longer natural. So Ulam later strengthened this to κ-*additivity*:

Definition 2.3. If μ is an Ulam measure on κ, we say that μ is κ-*additive* if for all disjoint families of size $< \kappa$, i.e. for all $\{A_\alpha : \alpha < \lambda\}$, with $\lambda < \kappa$, such that the A_α are pairwise disjoint, we have:

$$\mu\left(\bigcup_{\alpha<\lambda} A_\alpha\right) = \sum_{\alpha<\lambda} \mu(A_\alpha).$$

We wrote here "$\sum_{\alpha<\lambda}$", which looks a bit weird as we are potentially summing over a very large number of terms. But recall that μ is an Ulam measure, and so from our previous

[18] Indeed, this is the only "geometric" property, and so if we want just set-theoretic motivation it is natural to get rid of this assumption. All the other conditions make sense for arbitrary sets, but the translation invariance requires a notion of addition (+), which is not set-theoretic.

observation about the measure of disjoint sets, this is not really a sum, but a notation for:

$$\sum_{\alpha<\lambda}\mu(A_\alpha) := \begin{cases} 1 & \text{if and only if there is } \alpha<\lambda \text{ with } \mu(A_\alpha)=1; \\ 0 & \text{otherwise.} \end{cases}$$

Now we define:

Definition 2.4. A cardinal κ is called *measurable* if there is a κ-additive Ulam measure on κ.

Note that, just like before with σ-additivity and countable sets, μ being non-trivial along with being κ-additive implies that if $A \subset \kappa$ has cardinality $|A| < \kappa$, then $\mu(A) = 0$. So to have $\mu(A) = 1$, the set A must have cardinality κ. This will be an important observation later: in some sense, this means that κ being measurable means that there is a way of measuring the 'large' subsets of κ (namely, those with measure 1 are 'large', and those with measure 0 are 'small'), and having size $< \kappa$ means that you are automatically 'small'.

Before moving on, we introduce another terminology which turns out to be equivalent to the above definition of measurable cardinals: this is the terminology of *filters* and *ultrafilters*.

Definition 2.5. A family $\mathcal{F} \subset \mathcal{P}(\kappa)$ is called a *filter* (on κ) if it obeys:

- $\kappa \in \mathcal{F}$ and $\emptyset \notin \mathcal{F}$;
- $A, B \in \mathcal{F} \Longrightarrow A \cap B \in \mathcal{F}$;
- if $A \in \mathcal{F}$ and $B \supset A$, then $B \in \mathcal{F}$.

i.e. a filter contains κ, does not contain $\emptyset$, and is closed under finite intersections and "getting bigger" (i.e. supersets). In some sense, a filter contains 'large' subsets of κ.

Definition 2.6. We say that a filter $\mathcal{F}$ on κ is λ-*complete* if for all $\mu < \lambda$ and all collections $\{A_\alpha : \alpha < \mu\} \subset \mathcal{F}$ of sets in $\mathcal{F}$ of size μ we have $\bigcap_{\alpha<\mu} A_\alpha \in \mathcal{F}$.

i.e. if we strengthen the condition of being closed under finite intersections to being closed under intersections of size $< \lambda$. In particular, condition in a filter of being closed under finite intersections is equivalent to being $\aleph_0$-complete.

Definition 2.7. A filter $\mathcal{F}$ on κ is *non-trivial* if for all $\alpha \in \kappa$, we have $\{\alpha\} \notin \mathcal{F}$.

Definition 2.8. A filter $\mathcal{F}$ on κ is called an *ultrafilter* if for all $A \subset \kappa$, either $A \in \mathcal{F}$ or $\kappa \setminus A \in \mathcal{F}$.

The key observation which relates these filter definitions to our previous notion of Ulam measures is that κ-complete non-trivial ultrafilters (on κ) and κ-additive Ulam measures are the same thing, under the correspondence:

- Given an ultrafilter U, we get a κ-additive Ulam measure μ by setting: $\mu(A) = 1$ if and only if $A \in U$;
- Given a κ-additive Ulam measure μ, we get an ultrafilter on κ by setting: $A \in U$ if and only if $\mu(A) = 1$.

To check this, one just needs a bit of set algebra to check that the condition of κ-additivity and the condition of κ-completeness say the same thing in this setting. For example, if U is an ultrafilter, then κ-completeness is equivalent to (via taking complements): if for $\lambda < \kappa$, if $\{A_\alpha : \alpha < \lambda\}$ is such that $A_\alpha \notin U$ for all $\alpha < \lambda$, then $\bigcup_{\alpha<\lambda} A_\alpha \notin U$.

Let us now investigate the relationship between measurable cardinals and inaccessible cardinals, and in fact show that all measurable cardinals are inaccessible.

Theorem 2.1. *If κ is a measurable cardinal, then κ is regular.*

Proof. Let U be a κ-complete ultrafilter on κ. Then, as singletons are not in U, by κ-completeness we know that sets of size $< \kappa$ are not in U also.

So, suppose for contradiction that κ is not regular. Then, we can write $\kappa = \bigcup C$, where C is cofinal in κ and $|C| < \kappa$. What are the elements of C? They are elements of κ and thus ordinals in κ. But as κ is a cardinal, every ordinal in κ has size $< \kappa$. So, each element of C is a set of size $< \kappa$. Therefore, the elements of C cannot be in U, i.e. if $\alpha \in C$, then $|\alpha| < \kappa$ and so $\alpha \notin U$. But then κ is a union of size $|C| < \kappa$ of things not in U, and so by κ-completeness of U we would have $\kappa \notin U$; but this contradicts the definition of a filter. □

Theorem 2.2. *If κ is a measurable cardinal, then κ is a strong limit.*

Combining Theorem 2.1 and Theorem 2.2, we immediately get:

Corollary 2.1. *All measurable cardinals are inaccessible.*

Proof of Theorem 2.2. Suppose for contradiction that κ is a measurable cardinal which is not a strong limit. Then, by definition, there is $\lambda < \kappa$ such that $2^\lambda \geq \kappa$. But this means that we can inject κ into $\mathcal{P}(\lambda)$ or, more precisely, there is an injection $F : \kappa \to \{f : f : \lambda \to \{0,1\}\}$. Write $F_\alpha : \lambda \to \{0,1\}$ for the function $F(\alpha)$, where $\alpha \in \kappa$.

Now, fix any ordinal $\xi \in \lambda$. Then, clearly we have

$$\kappa = \{\alpha \in \kappa : F_\alpha(\xi) = 0\} \cup \{\alpha \in \kappa : F_\alpha(\xi) = 1\}$$

and this is a partition of κ into two disjoint sets. Thus, by the definition of an ultrafilter, at least one of these sets must lie in the ultrafilter U on κ (which exists and is κ-complete by assumption of κ being measurable). Let $b_\xi \in \{0,1\}$ be a value such that the set lies in U, i.e. b_ξ is so that $A_\xi := \{\alpha \in \kappa : F_\alpha(\xi) = b_\xi\} \in U$ (make a choice of value if needed).

We can do this for each $\xi \in \lambda$, and, as $\lambda < \kappa$, by κ-completeness of U we can take the intersection of these and it will still be in U, i.e. we know that

$$A := \bigcap_{\xi<\lambda} A_\xi \in U.$$

Now, we have

$$A = \bigcap_{\xi<\lambda} \{\alpha \in \kappa : F_\alpha(\xi) = b_\xi\} = \{\alpha : \text{for all } \xi < \lambda,\ F_\alpha(\xi) = b_\xi\}.$$

But, this set A can have at most 1 element, because it is a function which is determined at each $\xi \in \lambda$; indeed, here we are using injectivity of the map F, since if one has $\alpha, \alpha' \in A$, then $F_\alpha \equiv F_{\alpha'}$ are the same function $\lambda \to \{0,1\}$, i.e. $F(\alpha) = F(\alpha')$, and so $\alpha = \alpha'$ by injectivity of F. Thus, $|A| \leq 1$; but we know from non-triviality of U that U does not contain sets of size ≤ 1, and so we would need $A \notin U$, but this is a contradiction as we showed previously that $A \in U$. This contradiction proves the result. □

So we have now seen that measurable ⇒ inaccessible. Whether the reverse implication holds (which it does not) was an open question for quite some time.

Measurable cardinals will feature prominently in the second-half of this course (indeed, we will be exploring how the existence of these measures/ultrafilters has a huge impact on what is going on in the set-theoretic universe). For now, all we know is that they are large cardinals; indeed, proving that something is inaccessible is, at the moment, our main tool for showing that something is a large cardinal, as it implies that you cannot prove its existence in `ZFC` and that it has to be very big.

2.2. **Compactness.** In first-order logic, the Compactness Theorem is the statement that if a set of sentences Φ is finitely satisfiable (i.e. every finite subset is satisfiable), then Φ is satisfiable.

Compactness is both a blessing and a curse: whilst it is a very powerful tool to do things, it is the source of many of the limitative results about first-order logic: for example, we cannot characterise finite structures, countable structures, etc, and this can be shown through the compactness theorem. One idea to overcome this weakness was to expand language: Tarski's idea was to expand first-order logic to allow infinite formulas.

We are often frustrated, as sometimes we would like to say something like: $(x = 0) \vee (x = 1) \vee (x = 2) \vee \cdots$, i.e. have an infinite disjunction, but we can't write this as first-order logic demands that our formulas are finite. So, the idea is to just do it, i.e. think of the infinite disjunction as a formula. This is then called *infinitary logics*. Let us be more precise, and allow disjunctions and conjunctions a cardinal number of times, as well as quantifiers over a (potentially different) cardinal number of times.

Definition 2.9. Let λ, κ be cardinals. We then define $\mathcal{L}_{\kappa,\lambda}$ *languages* by:

(1) Allowing arbitrary families of variables[19];

(2) A set S of non-logical symbols (could be infinite or uncountable);

(3) $\wedge, \vee, \neg, \Rightarrow, \exists\forall$ exist as usual, with the usual semantics;

(4) we allow, for $\mu < \kappa$, conjunctions $\wedge_{\alpha<\mu}\phi_\alpha$ and disjunctions $\vee_{\alpha<\mu}\phi_\alpha$ of length μ;

(5) we allow, for $\mu < \lambda$, quantification over μ-many variables, i.e. we are allowed $(\exists^\mu x)\phi$ and $(\forall^\mu x)\phi$, where here x is a sequence of variables of length μ.

So, in $\mathcal{L}_{\kappa,\lambda}$ are are allowed conjunctions and disjunctions of length $< \kappa$ and quantifications to be of length $< \lambda$. In particular, $\mathcal{L}_{\omega,\omega}$ is just the language of first-order predicate calculus.

Terms are then constructed from S exactly as in first-order logic; in particular, atomic formulas are defined exactly as in first-order logic. To be precise, writing $M \vDash \wedge_{\alpha<\mu}\phi_\alpha$ means exactly

[19] i.e. instead of just countably many: so, if you like, take a proper class of variables so that we never run out.

that for all $\alpha < \mu$, $M \vDash \phi_\alpha$ (and similarly for disjunctions $\vee_{\alpha<\mu}\phi_\alpha$), and by $M \vDash (\exists^\nu x)\phi$ we mean exactly that there is a function $a : \nu \to M$ such that $M[x(\alpha) \mapsto a(\alpha)] \vDash \phi$, where by $M[x(\alpha) \mapsto a(\alpha)]$ we mean to interpret the variable $x(\alpha)$ by $a(\alpha)$.

We will be now interested in whether there is an analogue of the compactness theorem for $\mathcal{L}_{\kappa,\lambda}$ languages (and indeed, these analogues of compactness are what will give rise to the notions of *strongly compact* and *weakly compact* cardinals).

Note: This infinitary logic is far more powerful than first-order logic. Infinitary logic can express multiple features not expressible in first-order logic. For example:

$$\phi_{\text{fin}} := (\forall^\omega x)\left[\bigvee_{i\neq j}(x_i = x_j)\right];$$

this formula (which is a $\mathcal{F}_{\omega_1,\omega_1}$ formula; here, the variables are indexed by the natural numbers, i.e. $x_1, x_2, \dots$) describes structures which are finite, i.e. $M \vDash \phi_{\text{fin}}$ if and only if M is a finite structure. We know from the usual compactness theorem in first-order logic that ϕ_{fin} cannot be expressed in first-order logic (and hence this already shows that the usual formulation of the compactness theorem for first-order logic cannot be true for these infinitary languages in general).

Another example of this distinction is, if we write c_n for a constant symbol for each $n \in \mathbb{N}$, then we can put a bound on the size of the model by the formula

$$(\forall x)\left[\bigwedge_{n\in\mathbb{N}}(x = c_n)\right];$$

indeed, this describes models that are most countable, since if a model satisfies it, then each x must equal one of the c_n. Furthermore, if we also add $\bigvee_{n\neq m}(c_n \neq c_m)$, then this would say that the model is in bijection with $\mathbb{N}$ (i.e. it is countably infinite).

It is therefore somewhat naïve to think that the compactness theorem for first-order logic should hold for these more powerful languages, and therefore we define a different notion of compactness for them:

Definition 2.10. If Φ is a set of formulas in an $\mathcal{L}_{\kappa,\lambda}$-language and μ a cardinal, we say that Φ is *μ-satisfiable* if for each subset $\Phi_0 \subset \Phi$ with $|\Phi_0| < \mu$, then Φ_0 is satisfiable.

For example, being finitely satisfiable is equivalent to being $\aleph_0$-satisfiable in this terminology.

With this slightly more general notion, we can define the appropriate notion of compactness for infinitary languages: instead of demanding that everything that is finitely satisfiable is necessarily satisfiable, we demand for a cardinal κ that everything that is κ-satisfiable is satisfiable:

Definition 2.11. A cardinal κ is called *strongly compact* if for every $\mathcal{L}_{\kappa,\kappa}$-language L and every set Φ of L-sentences, the following holds: if Φ is κ-satisfiable, then Φ is satisfiable.

We can weaken this by adding an additional condition on the language L to get the notion of *weakly compact*:

Definition 2.12. A cardinal κ is called *weakly compact* if for every $\mathcal{L}_{\kappa,\kappa}$-language L with a set of non-logical symbols S of cardinality $|S| \leq \kappa$[20], and every set Φ of L-sentences, the following holds: if Φ is κ-satisfiable, then Φ is satisfiable.

Clearly strongly compact implies weakly compact.

If you look at these definitions, there is no a priori reason to believe that these are large cardinal notions. These are just definitions which are essentially obtained from generalising some notions from basic logic, and so it could very well be the case that these notions are not related at all to that of a large cardinal. Our main goal for the moment will be to see that these definitions in fact imply that κ must be inaccessible.

Let us start with a warm up and see that these notions do imply that κ is regular.

Theorem 2.3. *If κ is weakly compact, then κ is regular.*

Proof. Suppose not: then we can find κ which is weakly compact, yet $\kappa = \bigcup X$, where $|X| < \kappa$.

What can we do? We have essentially one fundamental idea of how we can use compactness that we already know from first-order compactness, and that is we can write down constant symbols and add something that can be satisfied as long as you only have very few formulas, but if you satisfy all of them at the same time we somehow get a contradiction. This is precisely the idea we will use here.

So take $\mathcal{L}_{\kappa,\kappa}$-language with constant symbols c_α, for $\alpha < \kappa$, and c[21]. The fundamental formula we use is then: $(c_\alpha \neq c)$ (and we have one such formula for each $\alpha < \kappa$); this just means that, if we are interpreting the constant symbols, then c will give rise to something which is different from all the c_α.

We now want to get t a contradiction by writing down a formula which expresses that everything is equal to a c_α. So. let us write the formula:

$$\phi := \bigvee_{\beta \in X} \bigvee_{\alpha < \beta} (c = c_\alpha)$$

i.e. ϕ says that there is a β in X and $\alpha < \beta$ such that $c = c_\alpha$. Note that $\bigvee_{\alpha<\beta}$ is a disjunction of size $\beta < \kappa$ (and this is $< \kappa$ as β is an element of κ), and $\bigvee_{\beta \in X}$ is a disjunction of size $|X| < \kappa$ (this is true by the contradiction assumption), and thus ϕ is an $\mathcal{L}_{\kappa,\kappa}$-formula.

Now let $\Phi := \{\phi\} \cup \{c_\alpha \neq c : \alpha < \kappa\}$. Clearly, Φ is κ-satisfiable (as if you have fewer than κ elements of Φ, then one of the c_α is missing and so you can interpret c as the one of the elements which is missing), but Φ is not satisfiable; but this contradicts κ being weakly compact; this contradiction completes the proof. □

Now that we know all weakly compact (and hence strongly compact) cardinals are regular, let us prove a key result, which links these compactness notions to the notions of filters and ultrafilters.

[20] Note that this is a $\leq \kappa$ and not a $< \kappa$.

[21] Note that this only uses κ-many non-logical symbols, which is important as we are only assuming that κ is *weakly* compact, and so this requires that the language has at most κ-many non-logical symbols.

Theorem 2.4 (Keisler–Tarski)**.** *Suppose that κ is strongly compact, and that $\mathcal{F}$ is a κ-complete filter on an arbitrary set X*[22]*. Then, $\mathcal{F}$ can be extended to a κ-complete ultrafilter.*

It should be noted that of course extending a filter to an ultrafilter is not the problem here (indeed, we can always extend a filter to an ultrafilter by just looking at all pairs of sets and their complements for which neither of the pair is in the filter, and just add one from each pair using `AC` to extend the filter to an ultrafilter). The crucial thing for the theorem is that you can do this whilst preserving κ-completeness.

We will see in the proof how this preservation of κ-completeness *really* requires the stronger version of our two compactness notions, as we will see how big the language needs to be in order to prove this.

Proof. Fix a set X and $\mathcal{F}$ a κ-complete filter on X. We will design a $\mathcal{L}_{\kappa,\kappa}$-language which is able to talk about these things.

Without loss of generality, using `AC`, we can assume that X is an ordinal, and hence elements of X directly correspond to subsets of X (via initial segments; we will see why having X an ordinal matters momentarily). If $A \subset X$, add a constant symbol c_A[23]. In addition, also add another constant symbol c.

Now, let L be the $\mathcal{L}_{\kappa,\kappa}$-language with non-logical symbols $\in$ and c_A (for $A \subset X$), and let L^* be the $\mathcal{L}_{\kappa,\kappa}$-language with non-logical symbols $\in$, c_A, and c. Now look at the following L-structure:

$$M := (\mathcal{P}(X), \in, \{A : A \subset X\})$$

i.e., in M we interpret our constant symbols c_A by A; this is where it matters that X is an ordinal, as only if elements of X are subsets of X can we express statements about membership in A in this structure.

Now let $\Phi := \mathrm{Th}_L(M)$, i.e. Φ is the $\mathcal{L}_{\kappa,\kappa}$-theory of this L-structure. In particular, we get that if $A \subset X$, then[24]

$$M \vDash (\forall x)(x \text{ is an ordinal} \to [x \in c_A \vee x \in c_{X\setminus A}])$$

and

$$M \vDash (\forall x)(x \in c_A \to x \text{ is an ordinal})$$

because c_A is interpreted by A and the complement $X \setminus A$ by $c_{X\setminus A}$ (and so these sentences are in Φ).

To the set of sentences Φ we add some extra formulas, forming:

$$\Phi^* := \Phi \cup \{(c \in c_A) : A \in \mathcal{F}\}.$$

We then have that Φ^* is κ-satisfiable: indeed, taking any subset of size $< \kappa$ in Φ^*, the term $(c \in c_A)$ only has to be true for $< \kappa$ many elements $A \in \mathcal{F}$ of the filter, say $(A_\alpha)_{\alpha<\lambda}$ for some $\lambda < \kappa$, but then by κ-completeness of $\mathcal{F}$ we know that $\cap_{\alpha<\lambda} A_\alpha \in \mathcal{F}$, and hence in

[22] We stress that there is not necessarily any relation between X and κ.

[23] Note that there are therefore $2^{|X|}$-many non-logical symbols to be added here, and so there is no way this argument could be done just using weak compactness, unless X is very small in size.

[24] Note that α is an ordinal if and only if $M \vDash$ (α is a transitive set totally ordered by $\in$).

particular this intersection is non-empty; thus, there is something in this intersection, and this something can be chosen as the interpretation of c.

Thus, by strong compactness of κ, we have that Φ^* is satisfiable. Hence, we have a model of Φ^*, say M^*, where c is interpreted as something, and it will lie in every single interpretation of the c_A (by definition of Φ^*). So define:

$$U := \{A \subset X : M^* \vDash (c \in c_A)\}.$$

By the above, we know that $M^* \vDash (c \text{ is an ordinal})$ (as M^* satisfied Φ and hence the two sentences listed above for M, in particular the second one). We now claim that U is a κ-complete ultrafilter on κ which extends $\mathcal{F}$.

Let us first check that U extends $\mathcal{F}$. This is obvious however, as we put in Φ^* all of the terms $(c \in c_A)$, for $A \in \mathcal{F}$, and so as $M^* \vDash \Phi^*$, all of these are true in U, and hence we immediately get $\mathcal{F} \subset U$.

To see that U is an ultrafilter, we already saw that M was a model of: $(\forall x)(x \text{ is an ordinal} \to [(x \in c_A) \vee (x \in c_{X\setminus A})])$, and as also $M^* \vDash (c \text{ is an ordinal})$, we get that in M^* this is also true for c, i.e. $M^* \vDash (c \in c_A) \vee (c \in c_{X\setminus A})$. But by definition of U, this would mean that precisely one of $A, X \setminus A$ is in U for each $A \subset X$, which means that U is an ultrafilter.

Finally we need to show that U is κ-complete; note that so far nothing has used infinitary languages, which is as expected as this is needed for the inheritance of κ-completeness. So let A_α (for $\alpha < \lambda$, where $\lambda < \kappa$) be such a family of sets of size λ in U, i.e. $M^* \vDash (c \in c_{A_\alpha})$. What we want to show is that c has to be in the intersection of these (as then this would imply that the intersection of the A_α lies in U, which is what we want for κ-completeness). So define:

$$A := \bigcap_{\alpha<\lambda} A_\alpha$$

and we want to show that $M^* \vDash (c \in c_A)$. How can we express this in infinitary logic? As A is an intersection, we see that A can be expressed as a size λ conjunction:

$$M \vDash (\forall x)(x \in c_A \leftrightarrow \bigvee_{\alpha<\lambda} (x \in c_{A_\alpha}));$$

this is clearly true in M, as each of these sets is interpreted as themselves, and this is exactly what it means to be an intersection, and so therefore (because the above is true $\forall x$, and hence the above is an L-sentence which lies in Φ)in M^* this is true for the new constant symbol c, i.e. we get

$$M^* \vDash (c \in c_A \leftrightarrow \bigvee_{\alpha<\lambda} (c \in c_{A_\alpha})).$$

But by above, our assumption was that $M^* \vDash (c \in c_{A_\alpha})$ for each $\alpha < \lambda$, and so this means that $M^* \vDash \bigvee_{\alpha<\lambda}(c \in c_{A_\alpha})$, and thus the above gives $M^* \vDash (c \in c_A)$. As explained, this implies that $A \in U$, proving that U is a κ-complete ultrafilter which extends $\mathcal{F}$. Thus, the proof is complete. $\square$

From Theorem 2.4, we can now readily prove that strongly compact implies measurable:

Corollary 2.2. *Any strongly compact cardinal is measurable.*

Proof. Recall that κ was a measurable cardinal if there was a non-trivial κ-complete ultrafilter on κ. So, by Theorem 2.4, we need to show that there is a κ-complete filter $\mathcal{F}$ on κ such that no extension of $\mathcal{F}$ to an ultrafilter can be trivial. So, look at:

$$\mathcal{F} := \{A \subset \kappa : |\kappa \setminus A| < \kappa\}.$$

By Theorem 2.3, we know that if κ is strongly compact then κ is regular, from which it follows that $\mathcal{F}$ is κ-complete (as otherwise, we would be able to write κ as a union of size $< \kappa$ of sets of size $< \kappa$).

Therefore, by Keisler–Tarski (Theorem 2.4), $\mathcal{F}$ can be extended to a κ-complete ultrafilter on κ, say $\mathcal{F}^*$. Moreover, we can see that $\mathcal{F}^*$ cannot be trivial, i.e. it cannot contain a 1-element set, because if it did, and say $A \in \mathcal{F}^*$ with $|A| = 1$, by definition we would have $\kappa \setminus A \in \mathcal{F} \subset \mathcal{F}^*$, and hence $\emptyset = A \cap (\kappa \setminus A) \in \mathcal{F}^*$, which is a contradiction to $\mathcal{F}^*$ being a filter. This contradiction shows that $\mathcal{F}^*$ is non-trivial, which completes the proof. $\square$

Now we have seen in total that: strongly compact $\Rightarrow$ measurable $\Rightarrow$ inaccessible. Our next goal will be to show that weakly compact $\Rightarrow$ inaccessible.

Theorem 2.5. *Any weakly compact cardinal is inaccessible.*

Proof. Let us revisit the proof that measurable $\Rightarrow$ inaccessible (namely, Theorem 2.1 and Theorem 2.2). What did we do there? Well, we already know that weakly compact $\Rightarrow$ regular (this is Theorem 2.3), and so we only need to prove that weakly compact $\Rightarrow$ strong limit. So, looking for a contradiction, let us suppose otherwise, i.e. κ is weakly compact, yet $2^\lambda \geq \kappa$ for some $\lambda < \kappa$.

In our language, we will now try to express that something exists that should be represented by a function $\lambda \to \{0, 1\} \equiv 2$, but it can't because we have put in all the formulas which say that it is not one of those functions.

So, we design an $\mathcal{L}_{\kappa,\kappa}$-language with constant symbols $c_\alpha, c_\alpha^0, c_\alpha^1$ for each $\alpha < \lambda$ (and note that the number of constant symbols we have here is $< \kappa$, and so this is OK for weakly compact).

Now let us try to express "c_α describes a function $\lambda \to 2$", where we think of c_α^0, c_α^1 as the values of the function at α, which could be 0 or 1. So, we need to write:

$$\phi^* := \bigwedge_{\alpha<\lambda} (c_\alpha^0 \neq c_\alpha^1) \tag{2.1}$$

i.e. ϕ^* is a formula saying that all of these are different, and we also want to say:

$$\phi^{**} := \bigwedge_{\alpha<\lambda} [(c_\alpha = c_\alpha^0) \vee (c_\alpha = c_\alpha^1)] \tag{2.2}$$

i.e. the value of the function at α is either 0 or 1 for each $\alpha \in \lambda$.

Let us write $\Phi := \phi^* \wedge \phi^{**}$. So, Φ really means that each c_α is really described by 0 or 1, i.e. exactly one of the c_α^0 or c_α^1 (this is why we include ϕ^*).

Let us now write, for each $f : \lambda \to 2$,

$$\phi_f := \bigvee_{\alpha<\lambda} (c_\alpha \neq c_\alpha^{f(\alpha)}) \tag{2.3}$$

i.e. ϕ_f describes if the function differs from f at some α. Let us write Φ^* for the theory of (2.1), (2.2), and (2.3).

Now, if something is a model of Φ, then in that model we have a function $\lambda \to 2$ described by whathever the c_α's are doing, i.e. if $M \vDash \Phi$, then we can define a function $g : \lambda \to 2$ by:

$$g(\alpha) := \begin{cases} 0 & \text{if } M \vDash (c_\alpha = c_\alpha^0); \\ 1 & \text{if } M \vDash (c_\alpha = c_\alpha^1). \end{cases}$$

So clearly Φ^* is inconsistent, as the additional formula (2.3) here says that, for every function f, it is <u>not</u> the case that c_α is described by f. But we have just seen that every model of Φ will give us a function that does this, and hence Φ^* is inconsistent.

The only thing left to do therefore is to show that Φ^* is κ-satisfiable, as we have just seen that it is not satisfiable, and so this would contradict κ being weakly compact.

So take any subset $\Phi_0 \subset \Phi^*$ with $|\Phi_0| < \kappa$. But then $|\Phi_0| < 2^\lambda$ by assumption, which means that Φ_0 cannot contain all of the additional formulas from (2.3) which made Φ^* inconsistent, i.e. there exists $f : \lambda \to 2$ with $\phi_f \notin \Phi_0$. But now we may interpret c_α by f, i.e. c_α gets the value $c_\alpha^{f(\alpha)}$, and this will then satisfy Φ_0. Hence Φ^* is κ-satisfiable, and therefore we have completed the proof. □

So now, to prove our original still of implications as listed at the start of this section, we just need to show: measurable $\Rightarrow$ weakly compact. For this, we will use *ultraproducts*.

2.3. **The Method of Ultraproducts.** Let U be an ultrafilter on κ[25] and let $M := (M_\alpha : \alpha < \kappa)$ be a sequence of first-order structures (it does not matter which language).

Now, we look at the collection of choice functions which pick an element from each of these structures, i.e. consider choice functions for M, that is functions $f : \kappa \to \bigcup_{\alpha<\kappa} M_\alpha$ such that $f(\alpha) \in M_\alpha$ for all $\alpha < \kappa$. We then define an equivalence relation on this collection of choice functions by:

$$f \sim_U g \iff \{\alpha : f(\alpha) = g(\alpha)\} \in U$$

i.e. f and g agree on a "large" set (where "large" means "is in the ultrafilter U"). This can readily be checked to be an equivalence relation on the set of such choice functions, and we define the *ultraproduct* to be the quotient of the choice functions modulo this equivalence relation, and we write $\mathrm{Ult}(M, U)$ for this quotient[26]. We also write $[f] := \{g : f \sim_U g\}$ for the equivalence class of f (which is an element of the ultraproduct).

[25]Technically we can do this for arbitrary sets, but as we will only need this for cardinals let us stick to this.

[26]So, an ultrafilter product is a quotient of a sequence of first-order structures by this relation determined by an ultrafilter.

In the special case where all the M_α agree, i.e. $M_\alpha = M_\beta$ for all $\alpha \neq \beta$, we call this an *ultrapower* instead of an ultraproduct (as we are essentially looking at combining many copies of the same structure, as opposed to combining different structures, so the 'product' is just a 'power'). We will abuse notation and then write M for this single model.

One can check that obvious definitions on the ultraproduct are well-defined and give a structure of the right type; for example, if R is a binary relation symbol on the M_α, then it also gives a relation on $\mathrm{Ult}(M, U)$ by:

$$[f]R[g] \iff \{\alpha < \kappa : M_\alpha \vDash f(\alpha)Rg(\alpha)\} \in U$$

i.e. $[f]R[g]$ if and only if the relation holds up to "U-equivalence". This means that, if there is a constant symbol, then it defines an object in the ultraproduct, and if there is a function it actually defines a function in the ultraproduct, and so on.

Now comes the important result regarding ultraproducts which we will need:

Theorem 2.6 (Łoś's Theorem[27])**.** *If ϕ is a formula in n variables, then:*

$$\mathrm{Ult}(M, U) \vDash \phi([f_1], \dots, [f_n]) \iff \{\alpha < \kappa : M_\alpha \vDash \phi(f_1(\alpha), \dots, f_n(\alpha))\} \in U.$$

This theorem says that, if we have any formula with n free variables, then we can describe precisely what it means for that formula to be true in the ultraproduct by the obvious generalisation of what we just did for the definition of (constant/function/relation) symbols.

Proof. We will not prove this here: the proof is essentially by induction on the complexity of ϕ and using the fact that U is an ultrafilter at the right moment; we remark that this result is not true if U is just a filter. □

Now, if we look at the claim in Łoś's theorem, we immediately see that it can be simplified when ϕ is a sentence (i.e. the case when $n = 0$/there are no free variables), because you see that a sentence is true in the ultraproduct if and only if it is true in an ultrafilter set of the structures that show up in the sequence, i.e. for ϕ a sentence,

$$\mathrm{Ult}(M, U) \vDash \phi \iff \{\alpha < \kappa : M_\alpha \vDash \phi\} \in U.$$

In particular, if $M_\alpha \vDash \phi$ for all α, then $\mathrm{Ult}(M, U) \vDash \phi$. For example, ultraproducts of models of ZFC are also models of ZFC, i.e. if we take $\phi = \mathtt{ZFC}$, then if $M_\alpha \vDash \mathtt{ZFC}$ for all α, then $\mathrm{Ult}(M, U) \vDash \mathtt{ZFC}$.

If $\mathrm{Ult}(M, U)$ is an ultra*power*, we can then define a map $j_U : M \to \mathrm{Ult}(M, U)$ by $j_u(m) := [c_m]$, where c_m is the constant (choice) function, i.e. $c_m(\alpha) = m$ for all $\alpha < \kappa$ (here, $m \in M$). Moreover, this map is an embedding, i.e. it is an injection, since if $m \neq n$ then $c_m \neq c_n$ (as they disagree at every $\alpha < \kappa$).

Definition 2.13. The map $j_U : M \to \mathrm{Ult}(M, U)$ defined above is called the *ultrapower embedding*.

[27] The name "Łoś" is pronounced like "wash".

If one looks at the statement of Łoś's theorem more closely, they realise that this map j_U is not just an embedding, but it is actually an *elementary* embedding, i.e., for a formula ϕ with n free variables, and $m_1, \ldots, m_n \in M$, we have:

$$M \vDash \phi(m_1, \ldots, m_n) \iff \mathrm{Ult}(M, U) \vDash \phi(j_U(m_1), \ldots, j_U(m_n)).$$

Indeed, to see this, note that the right-hand side is simply (by definition of j_U):

$$\mathrm{Ult}(M, U) \vDash \phi([c_{m_1}], \ldots, [c_{m_n}])$$

but by Łoś's theorem we know precisely when something like this is true: it is if and only if

$$\{\alpha < \kappa : M \vDash \phi(c_{m_1}(\alpha), \ldots, c_{m_n}(\alpha))\} \in U$$

is true, which, as the c_{m_i} are constant functions, is if and only if

$$\{\alpha < \kappa : M \vDash \phi(m_1, \ldots, m_n)\} \in U$$

is true, but as this condition here is independent of α, this set is either all of κ or $\emptyset$. Thus, to lie in U, this is if and only if it is all of κ, which is if and only if $M \vDash \phi(m_1, \ldots, m_n)$, which is what we wanted to show. □

We can therefore think of M as an elementary substructure of its ultrapower $\mathrm{Ult}(M, U)$ by identifying the elements of M with the equivalence classes of the constant functions, i.e. the image $j_U(M)$.

We want to now apply these observations to show that measurable ⇒ weakly compact.

Theorem 2.7. *Every measurable cardinal is weakly compact.*

Proof. For this, let us think about the size of the languages which only have κ-many non-logical symbols.

So, suppose κ is measurable. We know by Corollary 2.1 that κ is inaccessible.

Claim: If L is an $\mathcal{L}_{\kappa,\kappa}$-language with at most κ-many non-logical symbols, then $|L| \leq \kappa$.

Proof of Claim. The set of atomic formulas clearly has size at most κ, as it is precisely the number of non-logical symbols. So the only thing we need to do is show that the operations that we use to produce formulas in the $\mathcal{L}_{\kappa,\kappa}$-language preserve being of cardinality $\leq \kappa$.

Intuitively, we only need to show that the construction steps for $\mathcal{L}_{\kappa,\kappa}$-languages preserve "has cardinality $\leq \kappa$", i.e. if X is a set of formulas with $|X| \leq \kappa$, then the closure of X under the $\mathcal{L}_{\kappa,\kappa}$ has cardinality $\leq \kappa$.

Now, what is the cardinality if, say, we close under conjunctions of size $< \lambda$, for some $\lambda < \kappa$? A formula is now a sequence of length $< \kappa$ of things that were already present, and so in this case closing under $\bigwedge_{\alpha<\lambda} \phi_\alpha$, for $\lambda < \kappa$, will give cardinality $\kappa^{<\kappa} := \bigcup_{\lambda<\kappa} \kappa^\lambda$ (i.e. the cardinality of all sequences of size $< \kappa$).

But if κ is inaccessible, and thus in particular a regular strong limit, we have $\kappa^{<\kappa} = \kappa$. Indeed, if $\lambda < \kappa$, then $2^\lambda < \kappa$, as κ is a strong limit. But $\lambda^\lambda = 2^\lambda$, and so for any $\lambda, \mu < \kappa$, we have $\lambda^\mu < \kappa$. Now, if $f \in \kappa^{<\kappa}$, then there exists $\mu < \kappa$ with $f \in \kappa^\mu$. By regularity of κ, we know that f cannot be cofinal in κ^μ, and so there is some $\lambda < \kappa$ such that $\text{rank}(f) \subset \lambda$. Thus, $f \in \lambda^\mu$. This then implies that $\kappa^{<\kappa} \leq \kappa$; the other inequality is immediate as $\kappa \subset \kappa^{<\kappa}$.

Similarly, if you close under "$(\exists^\lambda x)\phi$ for some $\lambda < \kappa$, i.e. close under quantifiers, again you get sequences of length $< \kappa$ and so the cardinality is again $\kappa^{<\kappa} = \kappa$. This then completes the proof of the claim. ■

So we now know that if we take any $\mathcal{L}_{\kappa,\kappa}$-language with at most κ-many non-logical symbols, then I can think of any set of sentences in that language as being indexed by elements of κ, i.e. if Φ is any set of formulas in such an $\mathcal{L}_{\kappa,\kappa}$-language, we can write it as

$$\Phi = \{\phi_\alpha : \alpha < \kappa\}$$

by the claim above. We now need to show that if Φ is κ-satisfiable, then it is satisfiable (this would prove κ is weakly compact).

So suppose Φ is κ-satisfiable. Then, in particular every initial segment of it (in the sense of the above ordering of Φ, i.e. in the indexing $\Phi = \{\phi_\alpha : \alpha < \kappa\}$, will be satisfiable (as it is a set of formulas of size $< \kappa$). So define, for $\lambda < \kappa$, the initial segment determined by λ by:

$$\Phi_\lambda := \{\phi_\alpha : \alpha < \lambda\}.$$

Thus, we know by assumption of κ-satisfiability, Φ_λ is satisfiable. So, let $M_\lambda \vDash \Phi_\lambda$ be a model satisfying Φ_λ.

Then form $M = (M_\lambda : \lambda < \kappa)$, and form the ultraproduct $\text{Ult}(M, U)$, where U is a non-trivial, κ-complete, ultrafilter on κ (which exists by our assumption that κ is measurable). As usual, note that non-triviality and κ-completeness of U implies that U cannot contain any set of size $< \kappa$.

The claim is that $\text{Ult}(M, U) \vDash \Phi$. Indeed, to see this let $\phi \in \Phi$ (so $\phi \equiv \phi_\alpha$ for some $\alpha < \kappa$) be arbitrary. Then, what is

$$\{\lambda < \kappa : M_\lambda \vDash \phi_\alpha\}?$$

Well, look back at our choice of M_λ, if $\lambda > \alpha$, then $M_\lambda \vDash \phi_\alpha$, and so necessarily this set contains all $\lambda > \alpha$, i.e.

$$\{\lambda < \kappa : \lambda > \alpha\} \subset \{\lambda < \kappa : M_\lambda \vDash \phi_\alpha\}.$$

But $|\kappa \setminus \{\lambda < \kappa : \lambda > \alpha\}| = \alpha + 1$, which is $< \kappa$, and so this set cannot be in the ultrafilter U. Thus, as U is an ultrafilter, we know that $\{\lambda < \kappa : \lambda > \alpha\} \in U$. But then by the above containment, as U is a filter, we must have $\{\lambda < \kappa : M_\lambda \vDash \phi_\alpha\} \in U$. But then Łoś's theorem gives that $\text{Ult}(M, U) \vDash \phi_\alpha$. But as $\phi \equiv \phi_\alpha \in \Phi$ was arbitrary, this means that $\text{Ult}(M, U) \vDash \Phi$, and so Φ is satisfiable. Hence, we are done. □

3. A Non-Implication and Logical Hierarchies

We have now seen that:

$$\text{strongly compact} \implies \text{measurable} \implies \text{weakly compact} \implies \text{inaccessible}.$$

We will now show that inaccessible $\not\Longrightarrow$ weakly compact. We will do this by some sort of converse to Theorem 1.5, which said that if κ was inaccessible, then in V_κ we can always find elementary substructures. For weakly compact cardinals, we will prove that we can always find an elementary *super*structure, i.e. a bigger von Neumann class with V_κ being elementary in it.

This will give rise to an important technique known as *reflection*: for this, the main definition is the following (which, as said above, is 'dual' to the properties discussed for inaccessibles/worldly cardinals):

Definition 3.1. A cardinal κ is said to have the *Keisler extension property* (KEP) if there is a transitive set $X \supsetneq V_\kappa$ such that $(V_\kappa, \in) \preceq (X, \in)$, i.e. $(V_\kappa, \in)$ is an elementary substructure of $(X, \in)$.

We will use this to show that the least weakly compact cardinal *cannot* be the least inaccessible (and thus the least inaccessible cardinal is inaccessible, but not weakly compact, which proves the non-implication we want).

Our first step towards this is:

Theorem 3.1. *If κ is weakly compact, then κ has the KEP.*

Proof. Start by looking at the size of V_κ: observe that κ is inaccessible, and so by the proof of Hausdorff's theorem, this means that:

- if $x \in V_\kappa$, then $|x| < \kappa$;
- $|V_\kappa| = \kappa$.

The second property gives us an idea of how to formulate "being V_κ" in an $\mathcal{L}_{\kappa,\kappa}$-language. So, we define $\mathcal{L}_{\kappa,\kappa}$-languages L and L^* with additional constant symbols: (i) L has constant symbols c_x for each $x \in V_\kappa$; (ii) L^* has constant symbols the c_x ($x \in V_\kappa$) as well as an additional constant symbol c.

Observe that, by the second bullet point above, these languages have the additional property needed for weakly compact, namely that the size of the set of non-logical symbols is $\leq \kappa$ (indeed, it is equal to κ).

Now write:

$$\mathcal{V} := (V_\kappa, \in, \{x : x \in V_\kappa\})$$

i.e. we interpret each c_x by the correct set x in V_κ; this is an L-structure, and so we can take its theory; set $\Phi := \mathrm{Th}_L(\mathcal{V})$.

Now, as usual in compactness arguments, we now add statements saying that c is not equal to any of the other c_x. So set:

$$\Phi^* := \Phi \cup \{(c \neq c_x) : x \in V_\kappa\} \cup \{c \text{ is an ordinal}\}.$$

Now, if this was first-order logic, this would already be enough to produce an elementary equivalent model of the theory of V_κ where we would know there is an ordinal which is not equal to any of the c_x's. But here, when we use $\mathcal{L}_{\kappa,\kappa}$-languages, we can actually show slightly more. Observe that there are infinitary formulas that describe things in $\mathcal{V}$ that you cannot easily describe with a first-order formula, in particular:

$$\phi := (\forall^\omega x)\left(\bigvee_{i\in\mathbb{N}}[x_{i+1} \notin x_i]\right);$$

this is an $\mathcal{L}_{\omega_1,\omega_1}$-formula which says "for all sequences of variables, there is some natural number such that $x_{i+1} \notin x_i$" (i.e. this describes well-foundedness of $\in$).

So, as V_κ is well-founded, clearly $\mathcal{V} \vDash \phi$ (as $\phi \in \Phi$), i.e. $\mathcal{V}$ is well-founded. Since $\mathcal{V}$ is well-founded, anything we construct by compactness, because it will agree with the theory of $\mathcal{V}$, will also be well-founded, i.e. any model of $\Phi^* \supset \Phi$ will be well-founded.

Also, since κ is inaccessible (by Theorem 2.5), we know $\texttt{ZFC} \subset \Phi$, and in particular, the axiom of extensionality is in Φ. Hence, any model of $\Phi^* \supset \Phi$ will be extensional[28].

Now, clearly Φ^* is κ-satisfiable: indeed, if we have any collection of sentences $\Phi_0 \subset \Phi^*$ with $|\Phi_0| < \kappa$, we miss some term $(c \neq c_x)$, where $x \in V_\kappa$ (as $|V_\kappa| = \kappa$). In particular, as there are κ-many ordinals $< \kappa$ in V_κ, we miss some $(c \neq c_x)$ with x an ordinal, and so we take this x and make it the interpretation of c.

Hence, as κ is weakly compact, Φ^* is satisfiable; so take a model $M \vDash \Phi^*$. By our above observations, we know that M is extensional <u>and</u> well-founded (this is what we wouldn't get in first-order logic).

Thus, by Mostowski's Collapsing Theorem (see the *Logic and Set Theory* notes), there is a Mostowski collapse of M, i.e. there is a transitive set X such that $\mathcal{X} := (X, \in)$ is isomorphic to M with its membership (i.e. well-founded) relation, say E, i.e. $(X, \in) \cong (M, E)$.

We have now collapsed the structure M: M has things which were bigger than all ordinals occurring in V_κ (and so elements of κ), but by collapsing it we might have destroyed this. So, we first need to show that this Mostowski collapse is something containing V_κ itself.

So we first show: $V_\kappa \subset X$. For this we can again use the fact that we are using $\mathcal{L}_{\kappa,\kappa}$-languages, because the original structure obeys:

$$\mathcal{V} \vDash (\forall z)(z \in c_x \leftrightarrow \vee_{y\in x}(z = c_y))$$

(this is using the fact that $|x| < \kappa$ for $x \in V_\kappa$, as noted at the start of the proof, and so we are allowed to make this disjunction $\vee_{y\in x}$ as it has size $< \kappa$); this simply says that something is in x if and only if it is an interpretation of one of the constant symbols in x.

So, this formula is in Φ, and thus is true in M, and so by isomorphism it is true in X. Using this we can show, by a simple induction argument on the rank, that every element of V_κ is representing, in X, the constant symbol c_x, i.e. the interpretation of c_x in $\mathcal{X}$ is x, i.e. $c_x^{\mathcal{X}} = x$.

[28]Extensionality you would get with first-order compactness, but well-foundedness you would not get with first-order compactness; hence we are in a situation where we can do slightly more.

So, X is a set which contains all of V_κ, and the interpretation of c is an ordinal in X (by definition of Φ^*) which is bigger than all the ordinals in V_κ (by the extra formulas included in Φ^*), i.e. $c^{\mathcal{X}}$ is an ordinal, which is bigger than α for each $\alpha < \kappa$. In particular, $X \supsetneq V_\kappa$.

Therefore, $x \mapsto c_x^{\mathcal{X}} = x$ is the identity map, and by construction this is an elementary embedding (as we included the theory of the original model in Φ^*). Thus, we are done. □

Let us draw a picture of the situation we had in the above proof:

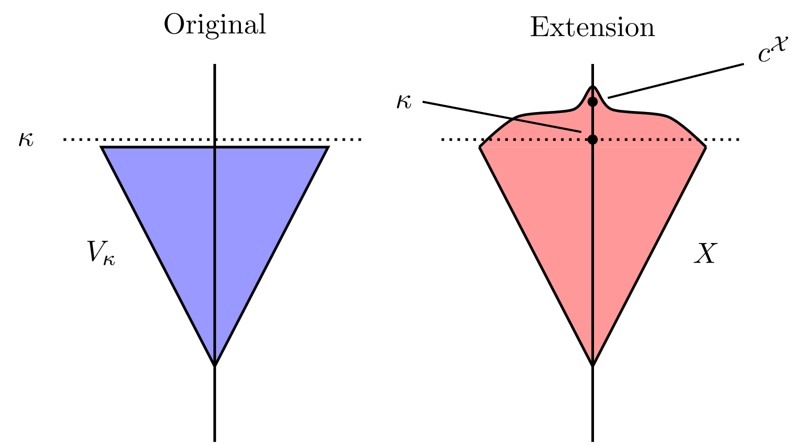

FIGURE 2. Illustration of the V_κ we start with and the extension X which contains V_κ as an elementary substructure. Note that X contains κ itself, and as it is transitive it must look something like shown (i.e. it is "closed downwards"). The construction gives that X must contain some interpretation of c which lies above κ and is an ordinal (the vertical line in each picture represents the ordinals in that model).

This picture now gives rise to the non-implication theorem we are after:

Theorem 3.2. *If κ is a weakly compact cardinal, then there is some cardinal $\lambda < \kappa$ with λ inaccessible.*

In particular, the following corollary is immediate:

Corollary 3.1. *The least inaccessible cardinal cannot be weakly compact. In particular, in general: inaccessible $\not\Rightarrow$ weakly compact (when, of course, an inaccessible cardinal exists).*

Proof of Theorem 3.2. Suppose that κ is a weakly compact cardinal. By Theorem 3.1 we know that κ has the KEP, and so we can find transitive $X \supsetneq V_\kappa$ such that $(V_\kappa, \in) \preceq (X, \in)$.

As we saw in the proof of Theorem 3.1, we have $\kappa \in X$. But, by absoluteness of inaccessibility for transitive models, i.e. inaccessibility is downwards absolute for transitive models of `ZFC` (see Proposition 1.4 and the surrounding discussion; we are just saying that if κ is inaccessible and $\kappa \in X$, then X will also think that κ is inaccessible), we have $(X, \in) \vDash$ "κ is inaccessible". In particular, this means that $(X, \in)$ is a model of `IC`, i.e. "there exists an inaccessible cardinal", i.e. $(X, \in) \vDash$ `IC`.

However, as $(V_\kappa, \in)$ and $(X, \in)$ are elementary equivalent (as being an elementary substructure implies elementary equivalent) and `IC` is a sentence, we necessarily have $(V_\kappa, \in) \vDash$ `IC`. But this means that there is some $\lambda < \kappa$ (as $\lambda \in V_\kappa$, and so $|\lambda| < \kappa$) such that

$$V_\kappa \vDash (\lambda \text{ is inaccessible}).$$

But now, since for models of the type V_κ inaccessibility is absolute (see Example Sheet 1, Q8), this tells us that λ really is inaccessible (and it is not just the case that V_κ thinks λ is inaccessible), and hence we have found $\lambda < \kappa$ which is inaccessible, which completes the proof. $\square$

Let us draw the picture again and see what is happening:

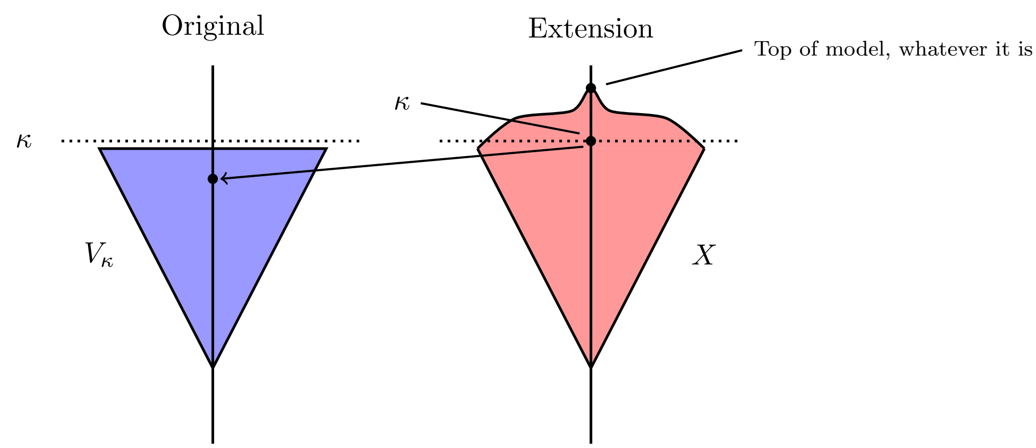

FIGURE 3. There is an inaccessible cardinal (namely, κ) which lies in the model X in the region $\geq \kappa$. This gets "reflected" by the elementary embedding to something in V_κ which is also inaccessible, and necessarily lies below the top of V_κ, i.e. below κ, which is what we wanted.

This phenomenon is called *reflection*: a property that holds in the taller universe, $\mathcal{X} = (X, \in)$ provided by KEP, is reflected downwards by the image of the elementary embedding to the shorter universe; note that it is not really the inaccessiblity of κ that makes this reflection work, as κ does not even exist in the smaller model, but it is the *elementary* that reflects the property downwards.

Let us briefly outline this reflection argument, giving a step-by-step analysis of the argument:

Step 1: Suppose we have some property that is true for κ itself, i.e. suppose that $\Phi(\kappa)$ holds. Now, of course we need to make sure that this property of κ still holds in the bigger model, $\mathcal{X} := (X, \in)$, so:

Step 2: Argue that $\mathcal{X} \vDash \Phi(\kappa)$; in the proof of Theorem 3.2, this was exactly the second paragraph.

Step 3: Therefore, $\mathcal{X} \vDash (\exists a)\Phi(a)$, which is now a sentence.

Step 4: (Reflection Step) Use the elementary substructure (and hence elementary equivalence) to get that $\mathcal{V} := (V_\kappa, \in) \vDash (\exists \alpha)\Phi(\alpha)$. Note that Step 3 was needed to pass to a sentence in order to do this, as κ is not in $\mathcal{V}$.

Now, we still need to get this "into the real world", and so we need to use some absoluteness argument to say that therefore there is some $\alpha < \kappa$ such that $\Phi(\alpha)$ holds, and it is not just that V_κ thinks that it holds, so:

Step 5: Argue that there is some $\alpha < \kappa$ such that $\Phi(\alpha)$ holds (this was the last paragraph in the proof of Theorem 3.2).

On Example Sheet 2, there are numerous exercises where you can see this reflection idea in action.

3.1. **Logical Hierarchies.** Now we will turn our attention to the hierarchy that we have for these large cardinals, as there are two completely different notions of "strength" that we have seen. One is the most obvious hierarchy, namely just be implication, and so for all the large cardinals we have seen are in an implication hierarchy (i.e. where one implies the other). But it seems from these arguments so far that there is something else going on: there is something about the size of these cardinals. So far, we have also actually seen that the cardinals are getting bigger and bigger. The question therefore is: what is the 'correct' notion of strength for a hierarchy of large cardinal axioms, i.e.

Q: How do we measure the logical strength of theories in general, or, more specifically to us, large cardinal axioms?

Our first naïve idea is *logical consequence*: for each theory, you can just look at its consequences. Indeed, if Φ is a set of sentences, then define:

$$C_\Phi := \{\phi : \Phi \vdash \phi\}$$

for its set of consequences. Then we can define (the *consequence hierarchy*):

$$\Phi \leq_0 \Psi \quad :\Longleftrightarrow \quad C_\Phi \subset C_\Psi.$$

This is nice for the axioms we have already seen. Indeed, we have already shown:

$$\mathtt{ZFC} \leq_0 \mathtt{ZFC} + \mathtt{IC} \leq_0 \mathtt{ZFC} + \mathtt{WC} \leq_0 \mathtt{ZFC} + \mathtt{MC} \leq_0 \mathtt{ZFC} + \mathtt{SC}$$

where these properties are: $\mathtt{WC}$ is $(\exists\kappa)(\kappa$ is weakly compact), $\mathtt{MC}$ is $(\exists\kappa)(\kappa$ is measurable), and lastly $\mathtt{SC}$ is $(\exists\kappa)(\kappa$ is strongly compact). Moreover, this is true in a very strong sense, as each large cardinal property above actually *implies* the one below, i.e. it does not just prove the existence of something with the property, but that cardinal itself also has the lower properties.

We can then say two theories are *logically equivalent*, written $\Phi \equiv_0 \Psi$, if (and only if) $(\Phi \leq_0 \Psi) \wedge (\Psi \leq_0 \Phi$, i.e. if $C_\Phi = C_\Psi$. Of course, we can then define $\Phi <_0 \Psi$ to be $(\Phi \leq_0 \Psi) \wedge (\Phi \not\equiv_0 \Psi)$.

We have then seen, from the non-provability of $\mathtt{IC}$ in $\mathtt{ZFC}$ (i.e. Corollary 1.1) and Corollary 3.1 that

$$\mathtt{ZFC} <_0 \mathtt{ZFC} + \mathtt{IC} <_0 \mathtt{ZFC} + \mathtt{WC}$$

(assuming the theories are consistent, of course).

This currently looks somewhat convincing that logical consequence might be a good way of looking at logical hierarchies. Unfortunately, this is never going to be a linear hierarchy: this hierarchy inherits the complexity from the fact it is really just a subset relation on a power set (which is inherently a partial order, not a linear/total order).

In particular, all of our theories we are dealing with, because of Gödel's incompleteness theorem, are incomplete theories, and if Φ is incomplete, then there is some ϕ such that $\Phi \nvdash \phi$ and $\Phi \nvdash \neg\phi$; but this means that $\Phi \cup \{\phi\}$ and $\Phi \cup \{\neg\phi\}$ are incomparable in $\leq_0$. In particular, we cannot compare $\mathtt{ZFC} + \mathtt{CH}$ and $\mathtt{ZFC} + \neg\mathtt{CH}$, where $\mathtt{CH}$ is the Continuum Hypothesis.

This is frustrating, and suggests that this might not be the right way to compare theories in general, and so it casts some doubt on $\leq_0$ as a good hierarchy.

Our second attempt is to use the notion of 'size'. The rough idea is that if we have two large cardinal properties, then one of them is 'stronger' than another if the cardinals are 'always bigger'.

So, let Φ be a *cardinal property*, i.e. if Φ obeys $(\forall x)(\Phi(x) \rightarrow x \text{ is a cardinal})$, and write $\Phi\mathtt{C} := (\exists x)\Phi(x)$. We now try to formalise the idea that "Φ-cardinals are bigger than Ψ-cardinals".

Obviously, the most naïve approach – that doesn't work – is to say that *every* Φ-cardinal is bigger than *every* Ψ-cardinal (of course this is not sensible or even possible, as we already know in many cases that a cardinal obeying a stronger cardinal property means it obeys the weaker one also). So instead, we are going to look at the *smallest* such cardinal obeying each property (this type of hierarchy is therefore only for cardinal axioms).

So, let ℓ_Φ be the smallest Φ-cardinal (if a Φ-cardinal exists, of course, then this will be well-defined – so in $\mathtt{ZFC} + \Phi\mathtt{C}$, ℓ_Φ is always defined and unique in each model). Then we define:

$$\Phi \leq_1 \Psi \quad :\Longleftrightarrow \quad \ell_\Phi \leq \ell_\Psi$$

(where this means that in any model where both exist). Similarly, we can define $\equiv_1$ and $<_1$. We already know from everything we have seen that[29]

$$\mathtt{ZFC} <_1 \mathtt{ZFC} + \mathtt{IC} <_1 \mathtt{ZFC} + \mathtt{WC} \leq \mathtt{ZFC} + \mathtt{MC} \leq_1 \mathtt{ZFC} + \mathtt{SC}$$

(we will improve on these latter two, and get them to be $<_1$, in the rest of the course).

So, this looks like a good notion of strength for large cardinal axioms, but we can easily come up with silly examples which don't fit our intuition for a good hierarchy.

For example, consider the property:

$$\Sigma(\kappa) \quad :\Longleftrightarrow \quad (\kappa \text{ is inaccessible}) \wedge (\exists\lambda)(\lambda \text{ is weakly compact})$$

i.e. Σ says that κ is inaccessible, and that there is some other cardinal (completely unrelated to κ potentially) which is weakly compact. Also write $\Sigma\mathtt{C}$ for $(\exists\kappa)\Sigma(\kappa)$ (we use Σ, i.e. the Greek letter 'S', to denote these 'silly' cardinals).

But clearly we have

$$\mathtt{ZFC} + \Sigma\mathtt{C} \iff \mathtt{ZFC} + \mathtt{WC};$$

[29]One should note that we can make $\mathtt{ZFC}$ itself into a "large cardinal axiom" (see Example Sheet 1, Q2) with the axiom of infinity acting as a sort of large cardinal axiom. Indeed, if we write $\mathtt{Inf}(\kappa) :=$ "κ is infinite", and $\mathtt{InfC}$ for $(\exists\kappa)(\kappa \text{ is infinite})$, then $\mathtt{ZFC} \Leftrightarrow \mathtt{ZFC} + \mathtt{InfC}$ (as there being an infinite cardinal follows from $\mathtt{ZFC}$).

indeed, the direction ($\Rightarrow$) follows from the definition of $\Sigma\texttt{C}$, as $\Sigma\texttt{C} \Rightarrow \texttt{WC}$, and the direction ($\Leftarrow$) follows from Theorem 3.2. In particular, the consequences of these two theories agree, i.e.

$$C_{\texttt{ZFC}+\Sigma\texttt{C}} = C_{\texttt{ZFC}+\texttt{WC}}$$

which implies that $\texttt{ZFC} + \Sigma\texttt{C} \equiv_0 \texttt{ZFC} + \texttt{WC}$. But of course, what is the size of the least Σ-cardinal? Well, $\Sigma(\kappa)$ just means for κ that it is inaccessible. However, both $\Sigma\texttt{C}$ and $\texttt{WC}$ holds in $\Sigma\texttt{C}$ holds, and thus by Theorem 3.2, we necessarily have $\ell_{\texttt{ZFC}+\Sigma\texttt{C}} = \ell_{\texttt{ZFC}+\texttt{IC}} < \ell_{\texttt{ZFC}+\texttt{WC}}$, and thus:

$$\texttt{ZFC} + \Sigma\texttt{C} <_1 \texttt{ZFC} + \texttt{WC}.$$

So, even though (in $\texttt{ZFC}$) the axioms are equivalent, the smallest such cardinal is much smaller.

As another silly example/property, consider the formulas

$$\phi(\kappa) := (\neg\texttt{WC}) \Rightarrow (\exists\alpha)(\aleph_\alpha \text{ is inaccessible and } \kappa = \aleph_{\alpha+\omega})$$

$$\phi'(\kappa) := (\texttt{WC}) \Rightarrow (\exists\alpha)(\aleph_\alpha \text{ is weakly compact and } \kappa = \aleph_{\alpha+\omega})$$

and set

$$\Sigma'(\kappa) := \phi(\kappa) \wedge \phi'(\kappa)$$

and as usual write $\Sigma'\texttt{C} := (\exists\kappa)\Sigma'(\kappa)$. Then what do we know about $\texttt{ZFC} + \Sigma'\texttt{C}$? Well, clearly if there are Σ'-cardinals, then there are inaccessible cardinals, and moreover if there are inaccessible cardinals then there are Σ'-cardinals[30]. Hence, we see that:

$$\texttt{ZFC} + \Sigma'\texttt{C} \iff \texttt{ZFC} + \texttt{IC}$$

and so as in the previous example, we would have $C_{\texttt{ZFC}+\Sigma'\texttt{C}} = C_{\texttt{ZFC}+\texttt{IC}}$, and hence $\texttt{ZFC}+\Sigma'\texttt{C} \equiv_0 \texttt{ZFC} + \texttt{IC}$.

But then note that under the assumption that both $\Sigma'\texttt{C}$ and $\texttt{WC}$ hold, then the smallest Σ'-cardinal is actually *bigger* than the smallest $\texttt{WC}$-cardinal (as any Σ'-cardinal is always the ω^{th}-successor of a weakly compact situation, when $\texttt{WC}$ holds. Thus, we have:

$$\texttt{ZFC} + \Sigma'\texttt{C} + \texttt{WC} \implies \ell_{\texttt{ZFC}+\texttt{WC}} < \ell_{\texttt{ZFC}+\Sigma prime\texttt{C}}$$

and so we get $\texttt{ZFC}+\texttt{WC} <_1 \texttt{ZFC}+\Sigma'\texttt{C}$ (despite, by the above observation on $\equiv_0$, that $\texttt{ZFC}+\Sigma'\texttt{C} <_0 \texttt{ZFC} + \texttt{WC}$).

These two examples illustrate that the 'size' hierarchy, $<_1$, is sometimes weird. The consequence hierarchy is much nicer as it reflects on 'strength' of axioms in a much more natural way than sizes, but on the other hand as consequences are covering everything it might be that a lot of things are not comparable.

What we want is a hierarchy which has the nice properties of the consequence hierarchy but somehow better captures the comparability. So to summarise, neither $\leq_0$ nor $\leq_1$ work nicely: we need something that behaves like $\leq_0$ without too much incomparability.

[30]Indeed, to see this note that if α is an inaccessible cardinal, consider its ω^{th} successor, i.e. take $\kappa := \aleph_{\alpha+\omega}$. Then if there are no weakly compact cardinals, then $\Sigma'(\kappa)$ is true, and if there are weakly compact cardinals, then if α' is a weakly compact cardinal, if we set κ' to be the ω^{th} successor of α', then $\Sigma'(\kappa')$ is true. Thus in either case, $\Sigma'\texttt{C}$ is true.

Remark: This phenomenon that $\leq_1$ does not always capture 'strength' well is usually in set theory called an *identity crisis*, following a paper by Magidor in 1974. This says roughly that if you have two large cardinal notions, $\Phi\mathtt{C}$ and $\Psi\mathtt{C}$, where $\Phi\mathtt{C}$ is clearly stronger than $\Psi\mathtt{C}$, yet there are models where the smallest cardinal of the first type is *equal* to the smallest cardinal of the second type (thus giving $\equiv_1$), i.e.

$$\Phi\mathtt{C} \wedge \Psi\mathtt{C} \wedge (\ell_{\Phi\mathtt{C}} = \ell_{\Psi\mathtt{C}})$$

is consistent in $\mathtt{ZFC}$. This phenomenon is of interest in research.

Now let us come to the hierarchy of most interest: this is called the *consistency strength hierarchy.*

To make things simpler for ourselves[31], here we will only do this for finite extensions of $\mathtt{ZFC}$. So, we focus on axioms systems which are of the form $\mathtt{ZFC}+A$, i.e. $\mathtt{ZFC}$ plus one additional axiom (which of course encompasses what we have been considering so far with large cardinal axioms). Now, consider the set of *consistency statements*:

$$\text{Cons} := \{\text{Cons}(\mathtt{ZFC}+A) : A \text{ is a formula}\}.$$

Note: All these formulas in Cons are formulas about natural numbers; indeed, modulo some coding or formulas, these are formulas of the type "for all natural numbers, that number is not a code for a proof where the last line is $(0=1)$, and all the axioms used in the proof are either $\mathtt{ZFC}$ axioms or A". So, modulo coding, the formulas $\text{Cons}(\mathtt{ZFC}+A)$ are all arithmetical formulas, where the quantifiers are bounded by $\mathbb{N}$; here they are Δ_0-formulas (as $\mathbb{N}$ is a set in our models of set theory and therefore is absolute between transitive models of set theory). Thus, all consistency statements are absolute between transitive models of set theory.

This observation helps us avoid some of the lack of comparability we had for general sets of formulas, as we had in $\leq_0$.

We can now define the *consistency strength hierarchy* for the axioms A by: $A \leq_{\text{Cons}} B$ if and only if the consequences of A restricted to the consistency statements are contained in the consequences of B restricted to the consistency statements, i.e.

$$A \leq_{\text{Cons}} B \quad :\Longleftrightarrow \quad \text{Cons} \cap C_{\mathtt{ZFC}+A} \subset \text{Cons} \cap C_{\mathtt{ZFC}+B}$$

i.e., we are using the consequence hierarchy, but we are now restricting this to just the consistency statements. We define $\equiv_{\text{Cons}}$ and $<_{\text{Cons}}$ in the usual manner.

Does this remove the problems we had before? It removes the silly problems as we are now essentially using the consequence hierarchy, so:

(1) The problems illustrated by Σ, Σ' go away, i.e. we cannot have identical consequences but strange relations with $\leq_1$. For example, for the same Σ, we had $C_{\mathtt{ZFC}+\Sigma\mathtt{C}} = C_{\mathtt{ZFC}+\mathtt{WC}}$, and hence $\text{Cons}\cap C_{\mathtt{ZFC}+\Sigma\mathtt{C}} = \text{Cons}\cap C_{\mathtt{ZFC}+\mathtt{WC}}$, and thus $\Sigma\mathtt{C} \equiv_{\text{Cons}} \mathtt{WC}$. Similarly, for Σ' we have $\Sigma'\mathtt{C} \equiv_{\text{Cons}} \mathtt{IC}$.

[31] To avoid talking about axiom schemes and what this means in the various models.

(2) Note that our proof of, say, $\mathtt{IC} <_0 \mathtt{WC}$ was by showing that there is a model of $\mathtt{ZFC}+\mathtt{IC}$ from $\mathtt{WC}$, i.e. $\mathtt{ZFC}+\mathtt{WC} \vdash \mathrm{Cons}(\mathtt{ZFC}+\mathtt{IC})$, and so we still remain the strict inequalities for $\leq_{\mathrm{Cons}}$, i.e. we still have

$$\mathtt{ZFC} <_{\mathrm{Cons}} \mathtt{ZFC}+\mathtt{IC} <_{\mathrm{Cons}} \mathtt{ZFC}+\mathtt{WC}$$

i.e. we did not lose anything in our hierarchy by just restricting to the consistency statements when forming $<_{\mathrm{Cons}}$ from $<_0$.

(3) What about the non-linearity issue? This is really where the power of using bounded formulas instead of all consequences lies. To illustrate this, let us look at what happens with $\mathtt{CH}$ and $\neg\mathtt{CH}$. We haven't said much about the $\neg\mathtt{CH}$ consistency proof, but recall our earlier mention of the $\mathtt{CH}$ consistency proof: this was done by Gödel in 1938, and this proof actually gives a *transitive* inner model where $\mathtt{CH}$ is true. This means that if we have a model of *anything*, then we are going to find a *transitive* submodel in which $\mathtt{CH}$ is true. But this means, as it is transitive and all consistency statements are absolute between the universe and transitive models, that it will have *exactly* the same consistency statements.

The same is true for the method of forcing, which was by Cohen to produces models of $\neg\mathtt{CH}$, where we also have that, if we have something, we will find a *transitive* model where $\neg\mathtt{CH}$ is true, and therefore it still retains all of the consistency statements.

So, this all means that our two methods which prove consistency of $\mathtt{CH}$ and consistency of $\neg\mathtt{CH}$ preserve Δ_0-statements, and therefore they preserve all consistency statements. So, any consistency statements that were true before you apply these methods will be true after you apply them, and neither of the two can prove any additional consistency statements to the ones proved in $\mathtt{ZFC}$ because if, say, $\mathtt{CH}$ proved something extra, then you can also produce a model of $\neg\mathtt{CH}$ where this extra statement is also true, and vice versa. So, this implies:

$$\mathrm{Cons} \cap C_{\mathtt{ZFC}} = \mathrm{Cons} \cap C_{\mathtt{ZFC}+\mathtt{CH}} = \mathrm{Cons} \cap C_{\mathtt{ZFC}+\neg\mathtt{CH}}$$

and so $\mathtt{ZFC} \equiv_{\mathrm{Cons}} \mathtt{ZFC}+\mathtt{CH} \equiv_{\mathrm{Cons}} \mathtt{ZFC}+\neg\mathtt{CH}$.

However, as a final remark: all is not well. Joel Hamkins recently published a preprint (on his webpage) on the *non*-linearity of cardinal consistency strength (see Theorem 2 there on *Rosser sentences*, which show that the consistency strength hierarchy is non-linear). So, this consistency strength hierarchy may get rid of the most silly examples, but there is still non-linearity present.

Now that we have our hierarchy, we will return and look at the remaining non-implications between our axioms. We will spend the most time on measurability; this was one of the big open questions in the 1930's which was essentially resolved using the technique of ultrapowers.

4. Measurable Cardinals and Elementary Embeddings

The aim of this section is to prove that: weakly compact $\not\Rightarrow$ measurable. We start in a slightly odd setting: in the end, we will like to do the present discussion in the theory $\texttt{ZFC}$ $+\texttt{MC}$, however we shall take a slightly stronger theory as it makes things easier to understand. Later we shall remove this extra assumption, and explain why it wasn't really necessary.

So, we first work out the assumption:

(A): *there exists κ which is measurable, and there exists $\lambda > \kappa$ which is inaccessible.*

This is obviously stronger (in the sense of $\leq_{\text{Cons}}$) than just assuming $\texttt{ZFC} + \texttt{MC}$, as it implies there is a model of $\texttt{ZFC} + \texttt{MC}$. Indeed, we know that $V_\lambda \vDash \texttt{ZFC} + \texttt{MC}$. Later we will remove the assumption that there is an additional, larger, inaccessible cardinal.

We let now fix U a κ-complete non-trivial ultrafilter on κ (which exists by measurability of κ). We can then form the ultrapower $\text{Ult}(V_\lambda, U)$; recall that this consists of (equivalence classes of) functions $f : \kappa \to V_\lambda$, where $f \sim_U g$ if and only if $\{\alpha \in \kappa : f(\alpha) = g(\alpha)\} \in U$.

Where do these functions actually live? If f is such a function, we realise that because λ is regular and $\kappa < \lambda$, this means that range(f) cannot be unbounded in V_λ (and thus must be bounded in V_λ). In particular, each such function f is actually an element of V_λ, i.e. $f \in V_\lambda$.

Of course, the equivalence classes are not elements of V_λ; this is because you can change a single value of f whilst remaining $\sim_U$-equivalent to f, and you can make this new value as high rank in V_λ as you like. But, all elements of the equivalence class are elements of V_λ, and so if we just take a representative for each equivalence class, then we can think of that as being a subset of V_λ, i.e. if we pick for each $\sim_U$-equivalence class a representative, and call the set of all such representatives X, then $X \subset V_\lambda$.

What is the $\in$-relation on this X? It is tempting to say "well, it is a subset of V_λ, so use the standard one on V_λ", but the $\in$-relation in the ultrapower is going to be very different from the standard $\in$-relation. So, X comes with the following $\in$, which we denote $\mathcal{E}$:

$$f \mathcal{E} g \quad :\Longleftrightarrow \quad \{\alpha \in \kappa : f(\alpha) \in g(\alpha)\} \in U$$

(this is natural, coming from the ultrapower). Clearly, $\mathcal{E} \neq \in$ (where $\in$ is the usual $\in$-relation on functions).

To summarise: so far we found in V_λ (remember that $\lambda > \kappa$ was inaccessible) this structure $(X, \mathcal{E})$, i.e. $X \subset V_\lambda$ and $\mathcal{E} \subset X \times X$, such that furthermore the map $V_\lambda \to X$ sending $x \mapsto c_x$, where c_x is the constant function with value x, is an elementary embedding from $(V_\lambda, \in)$ into $(X, \mathcal{E})$ (by Łoś's theorem).

It would be nice if we could find a subset of V_λ where the true $\in$-relation, i.e. $\in$, reflects $(X, \mathcal{E})$, i.e. we would like to find some *transitive* set $M \subset V_\lambda$ such that $(M, \in) \cong (X, \mathcal{E})$ via some isomorphism $\pi : X \to M$. Then, we would have an elementary embedding from V_λ into a transitive model, i.e. $V_\lambda \to M$, via the composition of $x \mapsto c_x$ and π.

To find this transitive M, we would like to use Mostowski's Collapsing Theorem, as we did previously (in the proof of Theorem 3.1). But to apply this, we need to know that $(X, \mathcal{E})$ is both well-founded and extensionable.

First note that $(X, \mathcal{E})$ being extensionable is no problem, as we know that λ inaccessible means that $V_\lambda \vDash \texttt{ZFC}$, and so in particular (as extensionality is an axiom of $\texttt{ZFC}$) $V_\lambda \vDash$ extensionality, and thus $(X, \mathcal{E}) \vDash$ extensionality. So, extensionality comes for free.

Well-foundedness however does not come for free, since in general ultrapowers will not be well-founded. Showing that $(X, \mathcal{E})$ is well-founded will use the fact that κ is measurable.

> **Aside:** (Triviality of Filters.) We defined filters to be non-trivial if they do not contain singletons: this is the right definition if we are talking about *ultra*filters; since most of the time we will be talking about ultrafilters, this is OK. However, if we are looking at general filters, this is not quite the right notions. There is a different notion of *freeness*:
>
> We say that a filter $\mathcal{F}$ is *free* if $\bigcap F = \emptyset$; if a filter is not free, we call it *fixed*.
>
> If U is an ultrafilter, U is free if and only if it is non-trivial. However, this equivalence is not in general true for filters (see Example Sheet 3).

Let us now prove that $(X, \mathcal{E})$ is well-founded:

> **Theorem 4.1.** $(X, \mathcal{E})$*, as constructed above, is well-founded.*

Proof. Suppose for contradiction that $(X, \mathcal{E})$ is not well-founded. Then, there is a sequence f_n of functions decreasing in $\mathcal{E}$, i.e. $[f_{n+1}]\mathcal{E}[f_n]$ for all $n \geq 1$. By definition of $\mathcal{E}$, this means that $A_n := \{\alpha \in \kappa : f_{n+1}(\alpha) = f_n(\alpha)\} \in U$ for each $n \geq 1$.

As there are only countably many $(A_n)_n$, by κ-completeness of the ultrafilter U, we have that $A := \bigcap_{n\geq 1}^{\infty} A_n \in U$, i.e. $A = \{\alpha \in \kappa : f_{n+1}(\alpha) \in f_n(\alpha) \text{ for all } n \geq 1\} \in U$. In particular, as $\emptyset \notin U$, we know $A \neq \emptyset$, and so pick $\alpha \in A$ to get $f_{n+1}(\alpha) \in f_n(\alpha)$ for all $n \geq 1$. But this is a decreasing sequence of $\in$ in V_λ, which is a contradiction to the wel-foundedness of V_λ. Hence, we are done. □

> **Note:** We did not need the full strength of κ-completeness of U in the above proof, but only $\aleph_1$-completeness. We only in fact proved that there are no infinite decreasing sequences with respect to $\mathcal{E}$: this is not quite well-foundedness, but it is equivalent to well-foundedness when we assume $\texttt{AC}$. This is not a problem however as we know that $(X, \mathcal{E}) \vDash \texttt{ZFC}$. In fact, $\aleph_1$-completeness of U is equivalent to the non-existence of infinite decreasing sequences (see Example Sheet 3).

So, with well-foundedness and extensionality being true in $(X, \mathcal{E})$, we can now apply the Mostowski Collapsing Theorem to collapse $(X, \mathcal{E})$ is a transitive set, $(M, \in)$, such that $(X, E) \cong (M, \in)$ via a bijection $\pi : X \to M$. Let us introduce a bit of notation for these models M:

Notation: If $f : \kappa \to V_\lambda$, we write $(f)_U := \pi([f]_U) \in M$ for the *Mostowski image* of f. We omit the subscript U and just write (f) when U is clear.

As π is an isomorphism, and hence an elementary embedding, we can now compose our elementary embeddings $V_\lambda \to X$ and $X \to M$ to get an elementary embedding $j \equiv j_U : V_\lambda \to M$; hence, $j(x) := (c_x)_U$ is the Mostowski collapse of the constant function c_x. Note that M is not defined to be the image of V_λ under j, and so we do not (and will not) necessarily have that j is a surjection.

Our next aim is to understand the set M and the embedding j. So, we now prove a serious of small propositions to get a better idea of what these look like. Let us start with the size of M.

Claim 1. $|M| = \lambda$.

Proof. Clearly we have $X \subset V_\lambda$ (thinking of X as the representatives, rather than the equivalence classes), but also $\{[c_\alpha] : \alpha < \lambda\} \subset X$, as there is a constant function $\kappa \to V_\lambda$ for each ordinal $\alpha < \lambda$ (and these are distinct for each $\alpha < \lambda$; hence we have an injective map $\lambda \to X$). Thus:

$$\lambda = |\{[c_\alpha] : \alpha < \lambda\}| \leq |X| \leq |V_\lambda| = \lambda$$

where the last equality comes from λ being inaccessible by assumption. Hence, $|X| = \lambda$. Hence, as $(X, \mathcal{E}) \cong (M, \in)$, we have $|M| = \lambda$. □

Claim 2. *If $x \in M$, then $|x| < \lambda$.*

Proof. Suppose $x \in M$. Then, by definition, there is $f : \kappa \to V_\lambda$ with $x = (f)_U$. As we have seen before, by regularity of λ, as $\kappa < \lambda$, we must have that the range of f is bounded below λ, i.e. we can find $\alpha < \lambda$ such that $f : \kappa \to V_\alpha$.

Now, if $y \in x$, then, by transitivity of M, $y \in M$, and so y is also represented by such a function, i.e. there exists $g : \kappa \to V_\lambda$ with $y = (g)_U$. Since $y \in x$, i.e. $(g)_U \in (f)_U$, by the isomorphism π this becomes, in X, that $[g]_{\sim_U} \in [f]_{\sim_U}$, and this we know is if and only if $\{\alpha \in \kappa : g(\alpha) \in f(\alpha)\} \in U$.

This means that, without loss of generality, we can think of g as having a range that goes into the elements of the range of f; hence we can without loss of generality assume that $g : \kappa \to V_\alpha$ as well.

So, we have seen that each $y \in x$ gives rise to a function $g : \kappa \to V_\alpha$ (and distinct y give rise to distinct g). So, to get an upper bound on $|x|$, it suffices to get an upper bound on the number of functions $\kappa \to V_\alpha$. But the total number of such functions is $|V_\alpha|^\kappa$, so we have $|x| \leq |V_\alpha|^\kappa$. But as λ is inaccessible and $\alpha < \lambda$, we know $|V_\alpha| < \lambda$ (this was Lemma 1.1), and therefore as λ is a strong limit, we have $|V_\alpha|^\kappa < \lambda$, and hence $|x| < \lambda$, as desired. □

We now know that all the sets $x \in M$ have few elements in comparison to λ. We can now prove:

Claim 3. $M \subset V_\lambda$.

Proof. By Claim 2, all elements of M have size $< \lambda$. Since M is transitive, this means that the transitive closure of each $x \in M$ has size $< \lambda$, i.e. $|\mathrm{TC}(x)| < \lambda$ for all $x \in M$, and hence $M \subset H_\lambda$ (where H_λ is the set of sets whose transitive closures have size $< \lambda$). But $H_\lambda = V_\lambda$, by Example Sheet 1, Q6 (using the fact that λ is inaccessible), and thus $M \subset V_\lambda$. □

So we have $M \subset V_\lambda$. But we can show that M contains *all* ordinals in V_λ:

Claim 4. $\mathrm{Ord} \cap M = \lambda$.

Proof. We already have seen that there is a well-ordered sequence of order-type λ in M: indeed, remember that $\{[c_\alpha] : \alpha < \lambda\} \subset X$, for where for each ordinal $\alpha < \lambda$, $c_\alpha : \kappa \to V_\lambda$ is the constant map, and so $\{(c_\alpha)_U : \alpha < \lambda\} \subset M$ is a set of ordinals of order-type λ (and note that these are just real ordinals, not just the ordinals in M, as c_α is the constant map and the Mostowski collapse therefore won't change them). So:

$$\lambda \leq \mathrm{Ord} \cap M \leq \mathrm{Ord} \cap V_\lambda = \lambda$$

i.e. $\mathrm{Ord} \cap M = \lambda$ (in the second inequality here we have used Claim 3). □

Remark: To expand on this point regarding ordinals not being changed: if $f : \kappa \to \lambda \equiv V_\lambda \cap \mathrm{Ord}$, then $(f)_U$ will be an ordinal. This is because in this situation, $\{\alpha < \kappa : f(\alpha) \text{ is an ordinal}\} = \kappa \in U$, and so by Łoś's theorem, $(X, \mathcal{E}) \vDash ([f] \text{ is an ordinal})$. Then, as $(X, \mathcal{E}) \cong (M, \in)$, we have $(M, \in) \vDash ((f)_U \text{ is an ordinal})$. But "being an ordinal" is absolute for transitive models of `ZFC`, and hence $(f)_U$ is an ordinal.

We can now draw a rough picture of our situation:

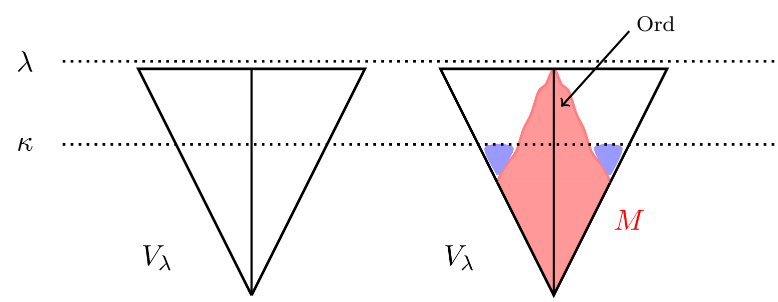

FIGURE 4. A depiction of the situation with all our current information: we start with V_λ, and via this Mostowski collapse, we construct M (shown in red). We know that $M \subset V_\lambda$, and that it is transitive (and hence is 'downwards closed', as shown). We also know that M contains all the ordinals in V_λ, and hence needs to contain the vertical line in V_λ, which represents the ordinals. Currently, this is the best picture we can draw. A natural question is then: does M really look like this? Somewhere, there is a line determined by κ: does M contain everything under κ? (i.e. need M contain the blue regions shown?) We will show that in fact we can draw a much better picture once we have proved more information regarding M.

We are now going to try and refine the above picture, and understand M better. The following claim tells us that the regions shaded blue in the above figure, i.e. the regions below the κ-line, do not actually exist, and that M contains everything below the κ-line in a very strong way:

Claim 5. $j_U|_{V_\kappa} = \mathrm{id}_{V_\kappa}$ *is the identity map on* V_κ.

Proof. We prove this by $\in$-induction on V_κ. So suppose $x \in V_\kappa$ is arbitrary, such that the claim is true for all $y \in x$, i.e. $j(y) = y$ for all $y \in x$. Note that, as j is an elementary embedding,

$$y \in x \iff (M, \in) \vDash (j(y) \in j(x))$$

but as $j(y) = y$ for all $y \in x$, this becomes:

$$y \in x \iff (M, \in) \vDash (y \in j(x)).$$

But $(y \in j(x))$ is an atomic formula, and atomic formulas preserve their truth values from transitive models (i.e. are absolute), and hence

$$(M, \in) \vDash (y \in j(x)) \iff y \in j(x).$$

So, we see that $y \in x$ implies $y \in j(x)$, and hence $x \subset j(x)$. Now we just need to show $j(x) \subset x$. So consider an element of $j(x) \equiv (c_x)_U$, i.e. a function $f : \kappa \to V_\lambda$ with $(f)_U \in (c_x)_U$. We know this is the case if and only if:

$$\{\alpha < \kappa : f(\alpha) \in c_x(\alpha)\} \in U$$

but as c_x is a constant function:

$$\{\alpha < \kappa : f(\alpha) \in c_x(\alpha)\} = \{\alpha < \kappa : f(\alpha) \in x\} = \bigcup_{y \in x} \{\alpha < \kappa : f(\alpha) = y\};$$

moreover, note that this last union is a union over a set of size $|x|$, and as κ is measurable (and thus inaccessible) and $x \in \kappa$ (by assumption), we have $|x| < \kappa$. To summarise: we have a union of size $< \kappa$ which lies in U. But now, as U is κ-complete, this means that at least one of these sets in the union must lie in U (this can be seen by taking complements), i.e. there exists $y \in x$ such that

$$\{\alpha < \kappa : f(\alpha) = y\} \in U.$$

But this would then imply that f is $\sim_U$-equivalent to the constant function c_y, i.e. $(f)_U = (c_y)_U$. But by induction, we know that $(c_y)_U \equiv j(y) = y$, and hence $(f)_U = y$. Hence we have shown that if $(f)_U \in j(x)$, then $(f)_U = y$ for some $y \in x$, i.e. $j(x) \subset x$. Combining we therefore have $x = j(x)$, which completes the proof (noting that, to complete the $\in$-induction, the result is clearly true for $\emptyset \in V_\kappa$). $\square$

We can now therefore draw an improved picture:

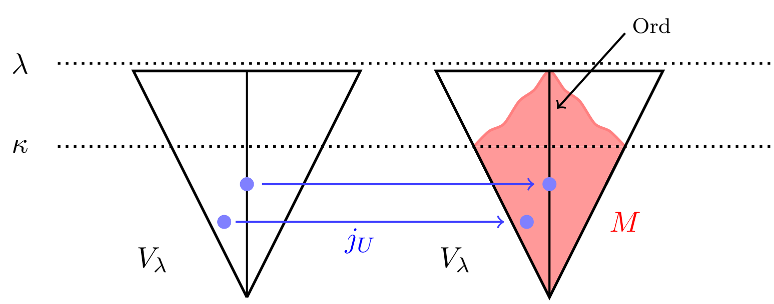

FIGURE 5. Our new picture of the situation is that M is transitive (so downwards-closed in the picture), contains all ordinals in V_λ (i.e. the vertical line), and contains everything below the line κ. We also know that j_U is the identity map below this line.

Note that we used κ-completeness in the proof of Claim 5 to get that the elements of $j(x)$ really are represented by *constant* functions, as we were able to use that, there at least, $|x| < \kappa$. So, if we take a set of size $\geq \kappa$, the above argument for showing that $j(x) \subset x$ is not going to work anymore.

So, if we wish to show that the elementary embedding j_U ($\equiv j$) is not the identity, we better try looking at the image of something which has size $\geq \kappa$. An obvious candidate for this would be κ itself. This leads us to the next claim:

Claim 6. *$j_U(\kappa) > \kappa$.*

This, j_U is the identity on V_κ, and this is exactly where it stops being the identity.

Proof. We have already seen before that, because κ is an ordinal, we must have that $j(\kappa)$ is an ordinal. Moreover, as elementary embeddings must be order-preserving on the ordinals, and as j is the identity map on V_κ, we must have that $j(\kappa) \geq \kappa$ (indeed, for all $\alpha < \kappa$ we have $j(\kappa) \geq j(\alpha) = \alpha$, by Claim 5).

So, if $j(\kappa) \neq \kappa$, we must have $j(\kappa) > \kappa$. We will therefore prove the claim by finding an ordinals that between all the $\alpha < \kappa$ and $j(\kappa)$, which then means that $\kappa \neq j(\kappa)$ and so $j(\kappa) > \kappa$.

Our candidate for this ordinal comes from the identity function, $\mathrm{id} : \kappa \to V_\lambda$. By the remark after Claim 4, we know that as id actually maps into $V_\lambda \cap \mathrm{Ord} = \lambda$, we know that $(\mathrm{id})_U$ is an ordinal.

Now, for any ordinal $\gamma < \kappa$, look at $A_\gamma := \{\alpha \in \kappa : \mathrm{id}(\alpha) > c_\gamma(\alpha)\}$. Note that as c_γ is the constant function and id the identity, we have $A_\gamma \equiv \{\alpha \in \kappa : \alpha > \gamma\} \equiv \kappa \setminus \gamma$; hence, by properties of the κ-complete ultrafilter U, we have $A_\gamma \in U$. This therefore tells us by our ordering in M (by $\in$) that $(\mathrm{id})_U > (c_\gamma)_U$. But $(c_\gamma)_U =: j(\gamma)$, and as $\gamma < \kappa$, by Claim 5 we have $j(\gamma) = \gamma$; combining, we therefore see that $(\mathrm{id})_U > \gamma$. Since $\gamma < \kappa$ was arbitrary, this shows that $(\mathrm{id}) \geq \kappa$[32].

But we can also look at: $\{\alpha \in \kappa : \mathrm{id}(\alpha) < c_\kappa(\alpha)\} = \{\alpha \in \kappa : \alpha < \kappa\} = \kappa \in U$, and thus this gives that $(\mathrm{id})_U < (c_\kappa)_U =: j(\kappa)$.

Combining, we therefore see that $\kappa \leq (\mathrm{id})_U < j(\kappa)$, i.e. $j(\kappa) > \kappa$, as claimed. □

In general, for λ inaccessible and $M \subset V_\lambda$ a transitive set, we say that an elementary embedding $j : V_\lambda \to M$ is *non-trivial* if $j \neq \mathrm{id}$. We can in fact show (see Example Sheet 3) that for each non-trivial embedding, there is an ordinal γ such that $j(\gamma) > \gamma$. The least such ordinal is then called the *critical point* of j, and denoted $\mathrm{crit}(j)$.

With this terminology, we can now formulate the central theorem regarding this topic:

Theorem 4.2 (Fundamental Theorem on Measurable Cardinals)**.** *Suppose λ is inaccessible and $\kappa < \lambda$. Then, the following are equivalent:*

(i) *κ is measurable;*

(ii) *There is a transitive set $M \subset V_\lambda$ and an elementary embedding $j : V_\lambda \to M$ with critical point κ.*

Proof. (i)⇒(ii): This follows from Claim's 1 – 6 above.

[32] i.e. if $(\mathrm{id})_U < \kappa$, then applying the above with $\gamma = (\mathrm{id})_U$ we would get $(\mathrm{id})_U > (\mathrm{id})_U$, a contradiction.

(ii)⇒(i): Start by observing that if we have such an elementary embedding with critical point κ, then it actually has the property we constructed before, namely it is the identity on V_κ; this follows by a standard induction on the rank[33].

Now define $U \subset \mathcal{P}(\kappa)$ by:

$$U := \{X \subset \kappa : \kappa \in j(X)\}.$$

Note that this definition makes sense: if $X \subset \kappa$, then as the map j is an elementary embedding, we have $j(X) \subset j(\kappa)$, and so as $\kappa \in j(\kappa)$ (as κ is the critical point), it makes sense to ask whether $\kappa \in j(X)$.

We claim that U is a κ-complete, non-trivial, ultrafilter on κ; this then implies that κ is measurable, as desired. Showing most of this will be just checking basic things arising from elementary embeddings.

Let us first prove U is non-trivial (in the sense of ultrafilters). If have $\alpha \in \kappa$ and look at the singleton set, $\{\alpha\} \subset \kappa$, what is $j(\{\alpha\})$? Well:

$$j(\{\alpha\}) = \{j(\alpha)\} = \{\alpha\}$$

where the first equality comes from the fact j is an elementary embedding, and the second from the fact $\alpha < \kappa$ and κ is the critical point of j, so $j|_{V_\kappa} = \mathrm{id}_{V_\kappa}$. So, for any $\alpha < \kappa$, we see that $\kappa \notin j(\{\alpha\})$, and so $\{\alpha\} \notin U$ for all $\alpha < \kappa$. So U is non-trivial.

Now let us see that U is a filter. We need to show that U is stable under intersections, supersets, and that $\emptyset \notin U$, yet $\kappa \in U$. Indeed, as j is an elementary embedding, we have (as intersections of sets is just conjugation):

- $j(X \cap Y) = j(X) \cap j(Y)$;
- $X \subset Y \implies j(X) \subset j(Y)$;
- $j(\emptyset) = \emptyset$ (as $\emptyset$ is defined by the x which satisfies $(\forall z)(z \notin x)$).

So, we immediately get from these that:

- $X, Y \in U \implies X \cap Y \in U$;
- $X \subset Y$ and $X \in U \implies Y \in U$;
- $\emptyset \notin U$.

To see that $\kappa \in U$, note that as $\kappa = \mathrm{crit}(j)$, we have $\kappa < j(\kappa)$, and so $\kappa \in j(\kappa)$ (as these are all ordinals), and so, as a set, $\kappa \in U$. All of this shows that U is a filter.

Now let us see that U is in fact an ultrafilter. Note that, again as j is an elementary embedding, $j(\kappa \setminus X) = j(\kappa) \setminus j(X)$. Hence, as $\kappa \in j(\kappa)$, we see that exactly one (and only one) of $j(X), j(\kappa \setminus X)$ will contains κ always. Hence, one of these sets will belong to U, which shows that U is an ultrafilter.

So, the last property of U which we need to show is that it is κ-complete. So fix some $\delta < \kappa$ and some sequence $\mathcal{A} = (A_\alpha : \alpha < \delta)$ of sets in U, i.e. $A_\alpha \in U$ for all $\alpha < \delta$. We want to then show $\bigcap_{\alpha<\delta} A_\alpha \in U$. Unravelling what this all means: we know that $\kappa \in j(A_\alpha)$ for all

[33]The rank is of course definable here.

$\alpha < \delta$, and we want to show that $\kappa \in j(\cap_{\alpha<\delta} A_\alpha)$. So, we need to understand the relationship between $j(\cap_{\alpha<\delta} A_\alpha)$ and $\cap_{\alpha<\delta} j(A_\alpha)$.

Now, we can think of $\mathcal{A}$ as a function $\delta \to \mathcal{P}(\kappa)$, i.e. it is a sequence of subsets of κ of length δ. This has the property that:

$$V_\lambda \vDash (\mathcal{A} \text{ is a sequence of subsets of } \kappa \text{ of length } \delta)$$

and so as j is an elementary embedding, we can push this forward to M to get:

$$M \vDash (j(\mathcal{A}) \text{ is a sequence of subsets of } j(\kappa) \text{ of length } j(\delta)).$$

Note that as $\delta < \kappa$ and $\kappa = \mathrm{crit}(j)$, we know that $j(\delta) = \delta$.

What do we know about the elements of this sequence $j(\mathcal{A})$? Well, similarly we have:

$$V_\lambda \vDash A_\alpha \text{ is the } \alpha^{\text{th}} \text{ element of } \mathcal{A}$$

and so again as j is an elementary embedding we have:

$$M \vDash j(A_\alpha) \text{ is the } j(\alpha)^{\text{th}} \text{ element of } j(\mathcal{A}).$$

But, as $\alpha < \delta < \kappa$, and $\kappa = \mathrm{crit}(j)$, we again know that $j(\alpha) = \alpha$. This means that we can identify $j(\mathcal{A})$ as the set of $j(A_\alpha)$'s, i.e.

$$j(\mathcal{A}) = (j(A_\alpha) : \alpha < \delta)$$

where we use round brackets, $(\bullet)$, to represent a sequence. But now what is $A := \cap_{\alpha<\delta} A_\alpha$? We have just shown that we can describe $j(A) = \cap_{\alpha<\delta} j(A_\alpha)$. Thus, if $\kappa \in j(A_\alpha)$ for each $\alpha < \delta$, then $\kappa \in j(A)$, i.e. $A \in U$, which is exactly what we wanted to show.

Thus, U is a κ-complete non-trivial ultrafilter on κ, which implies that κ is measurable, as desired. □

Remark: There is a remarkable similarity between this proof and the proof that strong compactness implies measurability (see Corollary 2.2).

Now let us return to our ultrapower construction of this elementary embedding j and transitive subset M under the assumption that κ was measurable, and analyse M even further. We start with some observations about cardinals in M.

Claim 7. *$j(\kappa)$ is measurable in M.*

Proof. Since $V_\lambda \vDash (\kappa \text{ is measurable})$, as j is an elementary embedding we get that $M \vDash (j(\kappa) \text{ is measurable})$. □

But this observation has a huge effect on the set theory that happens between κ and $j(\kappa)$, as we know that measurable implies that there are lots of inaccessible cardinals below. More precisely, if α is such that $\kappa < \alpha < j(\kappa)$, then there is some μ such that $\alpha < \mu < j(\kappa)$, and moreover $M \vDash (\mu \text{ is inaccessible})$ (compare with Theorem 1.4).

So, we started with $\kappa < \lambda$, where κ was measurable and λ was inaccessible, and we constructed M where M suddenly thinks that there are lots of inaccessibles between κ and λ (in M). This is our first main suggestion that maybe $M \neq V_\lambda$ (which is how we have always suggestively drawn our pictures): this is because we made no assumptions on the size of V_λ other than λ being inaccessible, and so if λ happens to be the *least* inaccessible above κ, then all of these μ found in M above which M thinks are inaccessible cannot be inaccessible in V_λ. So, at least in this particular case where λ is the least inaccessible above κ, we have $M \neq V_\lambda$ (we will later see that this is in fact true for any λ inaccessible, not just the least above κ).

Let us now look at what is preserved between V_λ and M. Since M is transitive in V_λ, we know that the notions of "cardinal", "regular", and "inaccessible" are downwards absolute (but not necessarily upwards absolute). But this means that something which is a cardinal, or a regular cardinal, or an inaccessible cardinal, in V_λ will still be a cardinal/regular/inaccessible in M. In particular, this is true for κ itself, which is inaccessible in V_λ, and thus $M \vDash (\kappa$ is inaccessible). This now allows us to perform a reflection argument, as we saw before.

Indeed, fix $\alpha < \kappa$. Then we know:

$$M \vDash \text{"there is } \mu \text{ such that } \alpha < \mu < j(\kappa) \text{ and } \mu \text{ is inaccessible"}.$$

But $\alpha < \kappa$, and so $j(\alpha) = \alpha$. Thus, we may be able to write this as:

$$M \vDash \text{"there is } \mu \text{ such that } j(\alpha) < \mu < j(\kappa) \text{ and } \mu \text{ is inaccessible"}.$$

So, as j is an elementary embedding, this becomes:

$$V_\lambda \vDash \text{"there is } \mu \text{ such that } \alpha < \mu < \kappa \text{ and } \mu \text{ is inaccessible"}.$$

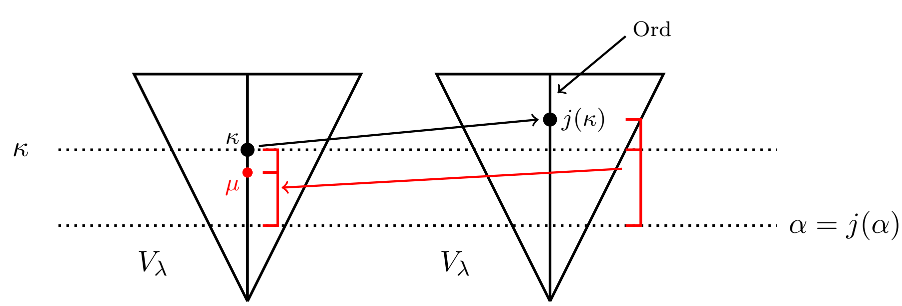

FIGURE 6. Performing the reflection argument. We find an interval in the image space, where the top and bottom cardinals are well-understood in the original V_λ, and there is something of interest between them. When we 'reflect' this back to the original space via the elementary embedding, we therefore find something of interest inbetween the cardinals in the original V_λ.

This is not a new result per se: we already know that measurable cardinals have unboundedly many inaccessible cardinals below, but the above is a new proof of this result which doesn't need weak compactness.

So, what do we know about these cardinals which lie between κ and $j(\kappa)$? Maybe, in the sense of V_λ, this is just ordinals of cardinality κ. This turns out to not be quite the case, as our next claim shows:

Claim 8. *$V_{\kappa+1} \subset M$.*

From this, we will see that there are cardinals in V_λ between κ and $j(\kappa)$.

Remark: We saw before that $V_\kappa \subset M$ in a very strong way, namely that $j|_{V_\kappa}$ was just the identity function on V_κ. We cannot expect this here, as we have $\kappa \in V_{\kappa+1}$ and we already know that $j(\kappa) \neq \kappa$ (by Claim 6), and so we know $j|_{V_{\kappa+1}} \neq \mathrm{id}|_{V_{\kappa+1}}$.

Proof. If $X \in V_{\kappa+1}$, then by definition we know that $X \subset V_\kappa$. We claim that in fact $X = j(X) \cap V_\kappa$; this would then imply that $X \in M$, as we know $V_\kappa = j(V_\kappa) \in M$, $j(X) \in M$, and M is transitive.

Let us first show $X \subset j(X) \cap V_\kappa$. If $x \in X$, then $x \in V_\kappa$ (as $X \subset V_\kappa$ and V_κ is transitive). So we just need to show $x \in j(X)$. But $x \in X$ and j being an elementary embedding gives that $j(x) \in j(X)$. But, as $j|_{V_\kappa} = \mathrm{id}|_{V_\kappa}$, we get $j(x) = x$, and so in fact we have $x = j(x) \in j(X)$. Combining, this shows $x \in j(X) \cap V_\kappa$; as this was for arbitrary $x \in X$, we therefore have $X \subset j(X) \cap V_\kappa$.

To see $j(X) \cap V_\kappa \subset X$, suppose $x \in j(X) \cap V_\kappa$. Then we have $x \in j(X)$ and $x \in V_\kappa$; this latter fact implies that $j(x) = x$ (just as above), and so with the former fact we have $j(x) = x \in j(X)$. By then as j is an elementary embedding, we again get that $j(x) \in j(X)$ implies that $x \in X$, which proves the claim. $\square$

What we do not currently know however is whether the model M thinks that the (cardinal) successor of κ is "what it should be". We can now prove this as a corollary of the above claim:

Corollary 4.1. *If κ^+ is the successor of κ in V_λ[34], then $M \models (\kappa^+$ is the successor of $\kappa)$.*

So, there are <u>no</u> cardinals between κ and κ^+ in M, i.e. if $\kappa < \alpha < \kappa^+$, then α is not a cardinal in M. In particular, as $j(\kappa)$ is a cardinal, we know that $j(\kappa)$ cannot lie between κ and κ^+, and so $j(\kappa) \geq \kappa^+$.

Proof. Take any ordinal α in M with $\kappa < \alpha < \kappa^+$ (i.e. M thinks this). Then (in reality, i.e. in original V_λ) there is a well-order of order-type α on κ (as $\alpha < \kappa^+$). But, this well-order is a subset of V_κ, and so an element of $V_{\kappa+1}$, and so by Claim 8, it must be an element of M. Therefore, M knows that α is isomorphic to something which lives on κ, and hence M knows that α cannot be a cardinal if $\alpha > \kappa$ (in M). Thus, $M \models$ "α is not a cardinal" for any such ordinal α. Thus, the first ordinal which M thinks is not in bijection with κ is κ^+, as desired. $\square$

We can now improve on our picture slightly, as we know that $V_{\kappa+1} \subset M$, and that κ^+ must lie somewhere between $\kappa + 1$ and $j(\kappa)$ (see below).

Now, you might think: "$j(\kappa)$ must be really big, because it is a measurable cardinal". But it is not a "real" measurable cardinal necessarily, it is only that M thinks it is a measurable cardinal. So, how big if $j(\kappa)$ really? We can actually give a bound on the size of $j(\kappa)$ in terms of the sizes of the cardinals in V_λ, since we know what the elements of $j(\kappa)$. Indeed:

[34] i.e. "in reality".

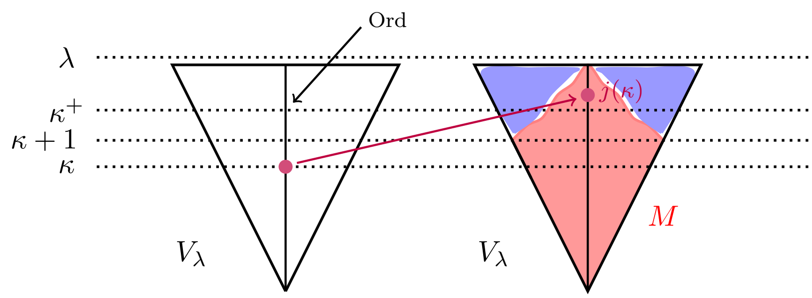

FIGURE 7. We now know that M looks something like shown. Of course, we do not yet know how large the blue region is (if it even exists: we currently do not know in general whether we could have $M = V_\lambda$ – we will see that this is never the case and so the blue region does exist in some form). All the dotted lines representing cardinal levels are the "true" values, i.e. those in V_λ.

Claim 9. *$|j(\kappa)| \leq 2^\kappa$, where here $|j(\kappa)|$ is the cardinality of $j(\kappa)$ as an ordinal in V_λ (i.e. "in reality").*

Proof. To see this, we just need to count the elements of $j(\kappa)$: every element of $j(\kappa)$ is represented by some function $f : \kappa \to V_\lambda$. So, if $(f)_U$ is an ordinal with $(f)_U \in j(\kappa)$, then as we have $j(\kappa) = (c_\kappa)_U$, for c_κ the constant function taking value κ, the relation $(f)_U \in (c_\kappa)_U$ actually mans that $\{\alpha \in \kappa : f(\alpha) = \kappa\} \in U$, and so without loss of generality we can assume (by choosing an appropriate representative of $(f)_U$), that $f : \kappa \to \kappa$. But this means that there are only at most $\kappa^\kappa = 2^\kappa$ many of these functions f, which proves the claim. □

But this means that in reality $j(\kappa)$ is quite small – indeed, despite M thinking it is a measurable cardinal, in reality it is not even a strong limit! So:

Corollary 4.2. *$j(\kappa)$ is not a strong limit cardinal (in V_λ).*

Proof. We know that, as ordinals, $j(\kappa) \geq \kappa^+$ (by the discussion after Claim 8), so in particular $j(\kappa) > \kappa$, yet $|j(\kappa)| \leq 2^\kappa$ by Claim 9. This means that $j(\kappa)$ cannot be a strong limit. □

Therefore, $j(\kappa)$ cannot be measurable (or inaccessible, even) in V_λ, as it is not a strong limit:

Corollary 4.3. *$j(\kappa)$ is not an inaccessible cardinal (in V_λ).*

So we now see a situation where $V_\lambda \vDash$ "$j(\kappa)$ is not inaccessible", yet $M \vDash$ "$j(\kappa)$ is inaccessible".

Using this, we can now show that, for any λ (and not just the smallest inaccessible above κ, as argued before) that $M \neq V_\lambda$. We will prove this by in fact producing a concrete example/witness of a set in $V_\lambda \setminus M$; in fact, this set is exactly the ultrafilter, U.

Theorem 4.3. *$U \notin M$. In particular, as $U \in V_{\kappa+2}$, we have $V_{\kappa+2} \not\subset M$.*

Using all of this, we now can illustrate the situation with the following (optimal) picture of the relationship between M and V_λ:

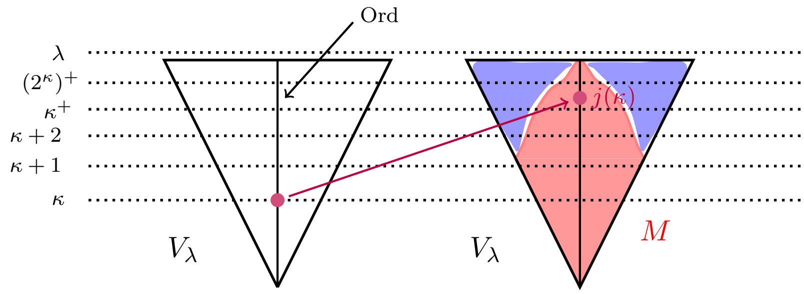

FIGURE 8. The final, optimal picture of M. We know that it contains everything below the $(\kappa+1)^{\text{th}}$ level, and but does not contain everything below the $(\kappa+2)^{\text{th}}$ level. Again, cardinal lines are those in V_λ, and the blue region represents the complement of M in V_λ.

Proof. The idea will be to use that fact that $V_\lambda \vDash (|j(\kappa)| \leq 2^\kappa)$ (as shown in Claim 9) and show that if $U \in M$, then the same is true in M, i.e. $M \vDash (|j(\kappa)| \leq 2^\kappa)$. This then gives a contradiction, as M thinks that $j(\kappa)$ is very large, namely that it is measurable.

So, let us suppose for contraction that $U \in M$ and see what additional information this would give us. In particular, we want to see if this gives us that M thinks there is a size bound on $j(\kappa)$.

The proof that $V_\lambda \vDash (|j(\kappa)| \leq 2^\kappa)$, as in Claim 9, was just by counting elements of $j(\kappa)$, as each of the elements of $j(\kappa)$ is represented by a (distinct) function $\kappa \to \kappa$. So, first of all note that if $f : \kappa \to \kappa$, then $f \in V_{\kappa+1}$[35], and so by Claim 8 we have $f \in M$ always. Therefore, the set of functions $\kappa \to \kappa$, namely κ^κ, obeys $\kappa^\kappa \in M$.

Our aim is now to construct a surjection, in M, from $\kappa^\kappa \to j(\kappa)$; if such a function exists in M, then we would get a size bound on $j(\kappa)$ in M. But of course, we know that, in V_λ, $j(\kappa) = \{(f)_U : f : \kappa \to \kappa\}$. So, our candidate surjection $\kappa^\kappa \to j(\kappa)$ is the function sending $f \mapsto (f)_U$; we need to show that this function exists in M.

Now, since we are assuming $U \in M$, we can actually look at the equivalence classes of $\sim_U$ in M: indeed, the equivalence relation $\sim_U$ on κ^κ defined by:

$$f \sim_U g \quad :\Longleftrightarrow \quad \{\alpha < \kappa : f(\alpha) = g(\alpha)\} \in U$$

can be formed in M, using the axiom of separation and using the fact that $U \in M$ (note that the axiom of separation holds in M as $M \vDash \texttt{ZFC}$).

So, $\sim_U \in M$, and thus the equivalence classes $[f]_{\sim_U}$ are elements of M, again by the axiom of separation. But now we have the Mostowski collapse (as this exists in all transitive models of $\texttt{ZFC}$), and so this defines uniquely the transitive set which is isomorphic to this, i.e. we can define, *in* M, the map $f \mapsto (f)_U$, where $(f)_U$ is the unique image of the Mostowski collapse of $[f]_{\sim_U}$.

This therefore shows that this function is actually a function in M (again, by the axiom of separation), and so arguing as before in the proof of Claim 9, we have $M \vDash (|j(\kappa)| \leq 2^\kappa)$. But this is a direct contradiction to $M \vDash (j(\kappa)$ is measurable$)$, as we know $M \vDash (j(\kappa) > \kappa)$ and so this in particular contradicts $j(\kappa)$ being a strong limit. □

[35] Indeed, ordered pairs of elements of κ have rank $< \kappa$, and therefore as f is a set of ordered pairs of elements of κ, we must have $f \subset V_\kappa$ and thus $f \in V_{\kappa+1}$.

Why have all just done all of this? All of this analysis was to try and understand if κ is still measurable in M. However, our main contender for witnessing that κ is measurable in M has just gone away: the reason κ was measurable in V_λ was because of the existence of the ultrafilter U in V_λ, but we have just seen that this ultrafilter does not exist in M. Of course, this does not necessarily mean that κ is not measurable in M, as there could be some other ultrafilter on κ in M.

To summarise: Theorem 4.3 clearly implies that κ cannot be measurable, in M, via the ultrafilter U, but it does not exclude the possibility that there is some other ultrafilter, U', on κ (in V_λ) such that U' 'survives' the ultrapower and is in M. In that case, κ would still be measurable in M.

Indeed, let us give this situation a name:

Definition 4.1. A cardinal κ is called *surviving* if it is measurable with an ultrapower embedding $j : V_\lambda \to M$ (as above) such that κ is still measurable in M, i.e. $M \vDash (\kappa \text{ is measurable})$.

The technique of reflection that we have seen several times already tells us that this notion must be strictly stronger than being measurable. Indeed, let us compare being surviving vs being measurable. So, if we suppose that κ is surviving (i.e. in V_λ), what this means is that, M thinks that not only that $j(\kappa)$ is measurable (this is Claim 7), but M also thinks that κ is measurable (by definition of surviving), and hence as $\kappa < j(\kappa)$ (by Claim 6), we have:

$$M \vDash \text{“there is } \mu < j(\kappa) \text{ such that } \mu \text{ is measurable”}$$

(indeed, this μ is just κ). But now we can reflect this back to the original V_λ, and we would get

$$V_\lambda \vDash \text{“there is } \mu < \kappa \text{ such that } \mu \text{ is measurable”}.$$

And so in particular, we see that the least measurable cardinal cannot be surviving.

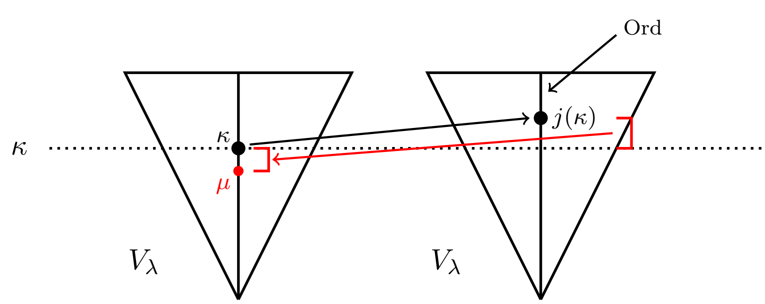

FIGURE 9. The reflection argument to find a smaller measurable cardinal when κ is surviving. The bracket which we reflect is determined by $j(\kappa)$ and κ.

Hence, we have shown:

Corollary 4.4. *If κ is surviving, then κ cannot be the least measurable cardinal. In particular, if κ is the least measurable cardinal, then it is not surviving, and so $M \vDash (\kappa$ is not measurable$)$.*

In particular:

- if κ is surviving, this implies that there are at least two inaccessible cardinals $\leq \kappa$, which is clearly stronger than there just being one $\leq \kappa$ (which is the case for the least measurable).
- measurability does not need to be preserved in the ultrapower embedding.

Again, using techniques we have seen before, we can push this even further: instead of just reflecting this one bracket (i.e. between $j(\kappa)$ and κ), we can reflect a bracket with 3 points, or, more generally, for all $\alpha < \kappa$, there is something between α and $j(\kappa)$ which is measurable in M (namely κ), and we can reflect this, using the fact that $j(\alpha) = \alpha$ (from Claim 5).

Indeed, fix any $\alpha < \kappa$. Then we know:

$$M \vDash (\exists\beta)(\alpha < \beta < j(\kappa) \text{ and } \beta \text{ is measurable})$$

(namely, $\beta = \kappa$; this follows as κ is surviving). Now, since $j(\alpha) = \alpha$ here, this can be rewritten as:

$$M \vDash (\exists\beta)(j(\alpha) < \beta < j(\kappa) \text{ and } \beta \text{ is measurable}).$$

This is now of the form where we can apply the fact that j is an elementary embedding, and reflect to get that:

$$V_\lambda \vDash (\exists\beta)(\alpha < \beta < \kappa \text{ and } \beta \text{ is measurable}).$$

Thus, as $\alpha < \kappa$ was arbitrary, this means that the set of measurable cardinals $< \kappa$ is unbounded (as we can simply inductively apply this to each new measurable cardinal we find). As κ is regular, this therefore implies that there are κ-many measurable cardinals below κ. So, we see that if κ is a surviving cardinal, then κ must be the κ^{th} measurable cardinal.

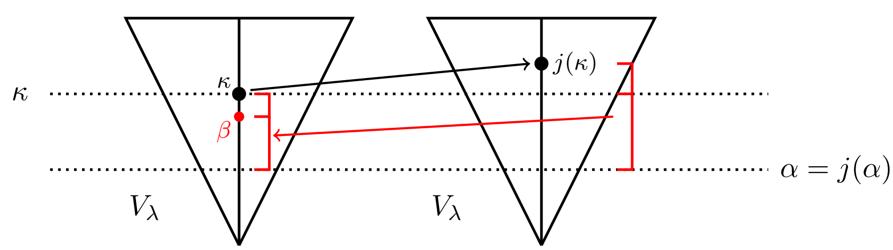

FIGURE 10. The improved reflection argument to find many smaller measurable cardinals when κ is surviving. The bracket which we reflect is determined by $j(\kappa)$, κ, and α ($< \kappa$).

So, we have now seen that measurability is not necessarily preserved in the ultrapower embedding. We have also seen being inaccessible *is* preserved in the ultrapower embedding. This is therefore a good situation to probe whether weakly compact implies measurability: weakly compact sits between these two notions, and so may or may not be preserved by the ultrapower embedding (and indeed, we only have to show that weakly compact is preserved for κ measurable, not in general). If it *is* preserved, then we would have a transitive model of `ZFC`, namely M, in which κ would be weakly compact (as this would be preserved from κ), but not measurable (if we took κ to be the smallest measurable cardinal in the universe/V_λ), which would therefore show that, in general, we have: weakly compact $\not\Rightarrow$ measurable, which is what we set out to prove from the start. Indeed, to be more precise, we would have this via the following reflection argument (which is the same as that for surviving cardinals just discussed):

Assume κ is measurable. Let $\alpha < \kappa$. Then, if $M \vDash (\kappa \text{ is weakly compact})$, we would have

$$M \vDash (\exists\beta)(j(\alpha) < \beta < j(\kappa) \text{ and } \beta \text{ is weakly compact})$$

which implies, reflecting using that j is an elementary embedding,

$$V_\lambda \vDash (\exists\beta)(\alpha < \beta < \kappa \text{ and } \beta \text{ is weakly compact}).$$

Hence this shows, in the same manner as for surviving cardinals, that κ would need to be the κ^{th} weakly compact cardinal in V_λ, i.e. we would have shown:

Corollary 4.5. *Assuming that $M \vDash$ "κ is weakly compact", then κ is the κ^{th} weakly compact cardinal (in the universe/V_λ).*

In particular, if α is any ordinal smaller than the least measurable cardinal, then the α^{th} weakly compact cardinal is not measurable. Hence, in general: weakly compact $\not\Rightarrow$ measurable.

So, the question now becomes: is κ weakly compact in M? Indeed, it is:

Theorem 4.4. *In the above setup, where $\kappa < \lambda$, where κ is measurable and λ is inaccessible, we have: $M \vDash$ "κ is weakly compact".*

Proof. Let L be an $\mathcal{L}_{\kappa,\kappa}$-language with at most κ-many non-logical symbols, say $S = \{s_\alpha : \alpha < \kappa\}$. Suppose that Φ is a set of L-formulas such that $M \vDash$ "Φ is κ-satisfiable", i.e. for any subset $\Phi_0 \subset \Phi$ of size $|\Phi_0| < \kappa$, there is a model (in M, say $N_{\Phi_0} \in M$) with $N_{\Phi_0} \vDash \Phi_0$.

To show that M thinks κ is weakly compact, we need to therefore prove that under these assumptions, $M \vDash$ "Φ is satisfiable". Now, as M is an inner model of V_λ (in particular, $M \subset V_\lambda$), we know that all these models N_{Φ_0} obey $N_{\Phi_0} \in V_\lambda$, and moreover as "$N_{\Phi_0} \vDash \Phi_0$" is a bounded formula (and therefore is absolute between V_λ and N_{Φ_0}), we also get that V_λ thinks that N_{Φ_0} satisfies Φ_0, and hence that $V_\lambda \vDash$ "Φ is κ-satisfiable".

This is now good, as in V_λ we know κ 0s measurable, and so in particular it is weakly compact in V_λ, and hence weak compactness gives:

$$V_\lambda \vDash \text{“}\Phi \text{ is satisfiable”}.$$

Hence, we find a structure $N \in V_\lambda$ with $N \vDash \Phi$.

There are two possible approaches from here to complete the proof – we will give one now and outline the other after completing the proof (the other proof will essentially boil down to proving a type of Löwenheim–Skolem theorem for $\mathcal{L}_{\kappa,\kappa}$-languages).

Our idea is to now use $j(N)$ and show that $M \vDash$ "$j(N) \vDash \Phi$", i.e. that M thinks $j(N)$ satisfies Φ (which will then prove that $M \vDash$ "Φ is satisfiable" and hence the result). This is a reasonable guess for a proof, but one needs to be careful.

We know that our language obeys $|L| \leq \kappa$ (and also $|\Phi| \leq \kappa$ – see the proof of Theorem 2.7), and so we can think of Φ as being indexed by κ, i.e. $\Phi = \{\phi_\alpha : \alpha < \kappa\}$.

Normally of course we do not really care about how we encode formulas in our set theory (as normally we do not do set theoretic operations with our formulas), but of course now we are: we are applying an elementary embedding to structures and formulas (and other things), and so we need to be very careful about how all of this is encoded. However, as κ is strongly inaccessible, we know that there is some way to ensure that all of these things are represented by elements of V_κ. So, let us assume that both $S, \Phi \subset V_\kappa$ (it doesn't matter how we encode them, as long as we encode them both as elements of V_κ, as if there are other things we don't really know what j is doing with them).

Let us now write down precisely what it means for V_λ to think N is a model of Φ:

$$V_\lambda \vDash (N \text{ is an } L\text{-structure and for all } \phi \in \Phi,\ N \vDash \phi)$$

(as we know, if ϕ is a single formula, then "$N \vDash \phi$" is a bounded formula and hence is nothing really complicated).

As j is an elementary embedding, this means that, mapping to M:

$$M \vDash (j(N) \text{ is a } j(L)\text{-structure and for all } \phi \in j(\Phi),\ j(N) \vDash \phi).$$

Now we see a problem: this is not a structure for our original language L, as now it is a $j(L)$-structure, and we are talking about elements in $j(\Phi)$, not Φ. So, we have two problems:

(a) $j(L)$ might not be L; indeed, it won't ever be L as L was an $\mathcal{L}_{\kappa,\kappa}$-language and so $j(L)$ is an $\mathcal{L}_{j(\kappa),j(\kappa)}$-language, and from Claim 6 we know $j(\kappa) > \kappa$, and hence $j(L)$ contains conjunctions/disjunctions/quantifiers of length $> \kappa$ (i.e. inbetween κ and $j(\kappa)$).

(b) $j(\Phi)$ might not be Φ.

Let us see how to deal with these two problems.

For (a), we want to understand what $j(L)$ is. We know it is an $\mathcal{L}_{j(\kappa),j(\kappa)}$-language, with non-logical symbols $j(S)$ (so, the situation is even worse, as this could contain non-logical symbols not in the original L). However, we know:

$$V_\lambda \vDash (\text{the } \alpha^{\text{th}} \text{ symbol in } S \text{ is } s_\alpha)$$

and thus as j is an elementary embedding, this implies

$$M \vDash (\text{the } j(\alpha)^{\text{th}} \text{ symbol in } j(S) \text{ is } j(s_\alpha)).$$

But note that as $\alpha < \kappa$, we know that $j(\alpha) = \alpha$ and as we assumed our coding was so that $s_\alpha \in V_\kappa$, we also know $j(s_\alpha) = s_\alpha$; both of these fact follow from that $j|_{V_\kappa} = \mathrm{id}|_{V_\kappa}$ (i.e. Claim 5). Hence, in fact this is just saying:

$$M \vDash (\text{the } \alpha^{\text{th}} \text{ symbol in } j(S) \text{ is } s_\alpha).$$

So, $j(S)$ has $j(\kappa)$-many symbols, but the first κ-many symbols are actually just the symbols of S. So we know $S \subset j(S)$. This is good as we can now just look at the reduct of $j(N)$, which just has symbols in S; i.e. we have an S-reduct of the $j(S)$-structure $j(N)$: call it $\overline{N}$.

Now let us address problem (b): what is $j(\Phi)$? By essentially the same argument as for (a) above, we can again say:

$$V_\lambda \vDash (\text{the } \alpha^{\text{th}} \text{ formula in } \Phi \text{ is } \phi_\alpha)$$

and hence as j is an elementary embedding, this gives for M

$$M \vDash (\text{the } j(\alpha)^{\text{th}} \text{ formula in } j(\Phi) \text{ is } j(\phi_\alpha))$$

but again, for $\alpha < \kappa$, we know that $j(\alpha) = \alpha$ and $j(\phi_\alpha) = \phi_\alpha$ (as $\phi_\alpha \in V_\kappa$ and $j|_{V_\kappa} = \mathrm{id}|_{V_\kappa}$), and hence again we get that the first κ-many formulas in $j(\Phi)$ are just Φ, and so $\Phi \subset j(\Phi)$.

Now we can put everything together to complete the proof: in M we have this structure $j(N)$, which is a model of $j(\Phi)$, and therefore as $\Phi \subset j(\Phi)$, we know that $j(N)$ is a model of Φ. Hence, its reduct to S, namely $\overline{N}$, is a structure in the right language which is a model of Φ, and therefore Φ is satisfiable in M. Hence

$$M \vDash (\overline{N} \vDash \Phi)$$

and so $M \vDash (\Phi \text{ is satisfiable})$. To conclude, this shows that $M \vDash (\kappa \text{ is weakly compact})$, which completes the proof. □

Before moving on, let us see an alternative proof of Theorem 4.4.

Alternative Proof of Theorem 4.4. Following the first three paragraphs of the original proof, we find a structure $N \in V_\lambda$ with $N \vDash \Phi$ (or, more precisely, $V_\lambda \vDash (N \vDash \Phi)$). We would like this N to be in M, as then we would be done (again, by absoluteness of $(N \vDash \Phi)$, M would know that Φ is satisfiable).

By why would $N \in M$? We would need to know what N is, and currently we don't understand N very well because it just came from weak compactness. So, can we understand N better? In particular, can we put any size bound on N? That would be related to the question (as we can pass to a different model): do we have some type of Löwenheim–Skolem theorem for these more general $\mathcal{L}_{\kappa,\kappa}$-languages?

Before discussing how one might prove a size bound, let us first consider if this is actually useful for our purposes. So, suppose we have the size bound $|N| \leq \kappa$ (or something which acts like N, with this size bound). This means that we can think of N as a structure on κ, so that $N = (\kappa, -)$, where the "$-$" represents the interpretations of the L-symbols in N. What are these interpretations? Well, they are relations (i.e. subsets of κ^n), functions (i.e. $\kappa^n \to \kappa$), and constant symbols (i.e. elements of κ).

But of course, all of these things (relations, functions, constants) would be in $V_{\kappa+1}$, and thus in M (by Claim 8)[36]. So, this means that if $|N| \leq \kappa$, then there is an isomorphic copy of N which lives in the part of V_λ (namely, $V_{\kappa+1}$) which is preserved in M, and therefore lives in M. But this would mean that M thinks Φ is satisfiable, which would complete the proof.

[36] Essentially, the structure is a κ-sequence of elements of $V_{\kappa+1}$, and so of M, and so we would like to see that M is closed under these κ-sequences.

So, our string of implications is: if $M \vDash (\Phi \text{ is } \kappa\text{-satisfiable})$, then $V_\lambda \vDash (\Phi \text{ is } \kappa\text{-satisfiable})$, which implies that $V_\lambda \vDash (\Phi \text{ is satisfiable})$, which implies that

$$V_\lambda \vDash (\Phi \text{ is satisfiable by a model of size} \leq \kappa) \tag{4.1}$$

which would then finally imply that $M \vDash (\Phi \text{ is satisfiable})$, as discussed above. So, the only implication we need to show to finish the proof is the third (i.e. the one giving (4.1)); note that this is essentially a type of Löwenheim–Skolem theorem for $\mathcal{L}_{\kappa,\kappa}$-languages, as we wish to change the size of the model. The proof of this can be found on Example Sheet 3, Q41). $\square$

To summarise, we have now seen how elementary embeddings and ultrapowers can be used to prove that weakly compact $\not\Rightarrow$ measurable. This was, however, under the assumption that there existed an inaccessible cardinal larger than our measurable cardinal – we will see how to address this point later.

Next, we look at various ways of strengthening measurability related to this idea.

5. Strengthenings of Measurability

Most of the large cardinal axioms that you find in Kanamori's table of large cardinal axioms are above measurable cardinals. One of the reasons we didn't see the definitions before was because most of their definitions are phrased in terms of the elementary embeddings which we have now seen for measurable cardinals. More precisely, many of the stronger large cardinal notions are defined using the elementary embedding we get from measurability, and demanding it to have some extra properties.

We will discuss 4 possible directions of strengthening measurability:

(1) Limit processes (this is exactly what we have done with other large cardinals before, and is not really related to the elementary embedding, and indeed we will see that it doesn't really lead to any new ideas);

(2) Survival (which we also saw earlier in some form);

(3) Strength;

(4) Supercompactness.

Let us work our way through this list.

5.1. **Limit Processes.**

This is a rather generic definition we have seen before for other large cardinals: we just require the thing to be a limit of smaller cardinals of the same type, and this is a somewhat stronger notion. Precisely:

Definition 5.1. We define inductively on ordinals α the notion of a cardinal κ being *α-measurable* via:

- κ is 0-*measurable* if κ is measurable;
- κ is called $(\alpha+1)$-*measurable* if it is α-measurable and $\{\mu : \kappa : \mu \text{ is } \alpha\text{-measurable}\}$ is unbounded in κ;
- for λ a non-zero limit ordinal, we say that κ is λ-*measurable* if κ is α-measurable for all $\alpha < \lambda$.

You could of course make similar definitions for, e.g., inaccessible cardinals, to get an increasing hierarchy of stronger and stronger axioms.

What we saw after the discussion of Corollary 4.4 (for surviving cardinals) is that if measurability of κ survives in the ultrapower, then a standard reflection argument gives that $\{\mu < \kappa : \mu \text{ is measurable}\}$ is unbounded in κ, i.e.

$$\kappa \text{ surviving}^{37} \implies \kappa \text{ is 1-measurable.}$$

But now we can repeat this: by exactly the same reflection argument used to prove this implication (as now κ is 1-measurable in M, and this gets reflected back), we get, for all $\alpha < \kappa$:

$$M \vDash (\exists\beta)(\alpha < \beta < j(\kappa) \text{ and } \beta \text{ is 1-measurable})$$

[37] Where this is how we defined it in Section 4.

and hence as j is an elementary embedding and $j(\alpha) = \alpha$ for $\alpha < \kappa$,

$$V_\lambda \vDash (\exists \beta)(\alpha < \beta < \kappa \text{ and } \beta \text{ is 1-measurable})$$

and so in particular:

$$\kappa \text{ surviving} \implies \kappa \text{ is 2-measurable.}$$

Repeating this by transfinite induction, we see that for all $\alpha < \kappa$, κ surviving implies that κ is α-measurable. Hence:

$$\kappa \text{ surviving} \implies \kappa \text{ is } \kappa\text{-measurable.}$$

5.2. **Survival.** We have already said in Section 4 that survival really means that there are two different ultrafilters living on κ; indeed, one of them is the initial one, and the second one comes from the ultrapower which survives in the ultrapower embedding (giving measurability of κ in the M). Hence, survival is really a relation on ultrafilters of κ. One could define a relation between ultrafilters on κ via:

$$U < V \quad :\iff \quad V \in M_U := \pi[\mathrm{Ult}(V_\lambda, U)]$$

i.e. $U < V$ (read "V survives U") is V is an element of the Mostowski collapse of the ultrapower with U.[38] So, if we have two ultrafilters, we can express whenever one of them survives the other. Hence, we could now define "surviving", as seen in Section 4, to be: κ is surviving if and only if there are two ultrafilters on κ, and one survives the other.

For further analysis, let us look at what this surviving really is, i.e. what does it mean that $U < V$? It means that in the ultrapower with respect to U there is an object that is V. Hence, as elements of $M_U \equiv M$ are determined by functions, it means there is a function $g : \kappa \to V_\lambda$ such that $(g)_U = V$.

What can we say about g? Well, M_U thinks that $(g)_U = V$ is a κ-complete ultrafilter on κ. We also know that $\kappa \in V_{\kappa+1} \subset M_U$ (by Claim 7), and hence there exists some function $f : \kappa \to V_\lambda$ (in fact, $f : \kappa \to \kappa$) with $\kappa = (f)_U$. In this language, we therefore have:

$$M_U \vDash \text{"}(g)_U \text{ is a } (f)_U\text{-complete ultrafilter on } (f)_U\text{"}$$

and hence, as M_U is an ultrapower, Łoś's theorem gives that this is if and only if:

$$\{\alpha < \kappa : g(\alpha) \text{ is an } f(\alpha)\text{-complete ultrafilter on } f(\alpha)\} \in U.$$

Note that $f(\alpha) < \kappa$, and so as $g(\alpha)$ is an ultrafilter on $f(\alpha)$, it lives two ranks above $f(\alpha)$ in the von Neumann hierarchy, and so in particular $g(\alpha) \in V_\kappa$ as well. But, as $f : \kappa \to \kappa$ is a function, we know $f \in V_{\kappa+1}$, and hence as we have just seen that $g : \kappa \to V_\kappa$, we also have (as a function) $g \in V_{\kappa+1}$.

[38]One may ask if this $<$ really is a strict order, i.e. whether it is irreflexive and transitive. Irreflexivity follows from the fact $U \notin M_U$, which we say in Theorem 4.3 (which was showing $V_{\kappa+2} \not\subset M$). Transitivity is not obvious, and we won't discuss in this course.

In particular, we see that, in general, if you have any inner model M in V_λ which obeys $V_{\kappa+1} \subset M$ and that $U, V \in M$ are ultrafilters on κ, then:

$$M \vDash (U < V) \iff V_\lambda \vDash (U < V). \tag{5.1}$$

Hence, survival of ultrafilters is a property which is preserved as long as $V_{\kappa+1}$ is contained in that model (which is the true for the situation regarding measurable cardinals).

Now we have this notion of surviving ultrafilters, we can use this to define a rank function:

Definition 5.2. A cardinal κ has *Mitchell rank* $\geq n$, and write $o(\kappa) \geq n$, if there are κ-complete non-trivial ultrafilters $U_0, U_1, \dots, U_n$ on κ such that $U_0 < U_1 < \cdots < U_n$[39].

We say that κ has *Mitchell rank* n, and write $o(\kappa) = n$, if it has Mitchell rank $\geq n$ but does not have Mitchell rank $\geq n+1$.

Note that we can of course define Mitchell rank α, for any ordinal α, by continuing this definition transfinitely in the obvious manner. Of course, $o(\kappa) \geq 0$ is equivalent to κ being measurable, and what we defined as 'surviving' in Section 4 is equivalent to having $o(\kappa) \geq 1$.

Let us see that this really is a hierarchy, namely that $o(\kappa) \geq n+1$ is strictly stronger than $o(\kappa) \geq n$. We will just do this in the case $n = 1$, and the rest follows from induction.

So suppose that $o(\kappa) \geq 2$. Then, we can find κ-complete non-trivial ultrafilters $U_0 < U_1 < U_2$ on κ. Let us look at what happens in the ultrapower by U_0: let M_0 be this ultrapower. Then, as $U_0 < U_1$ and $U_0 < U_2$, we have that U_1, U_2 survive in M_0, and, from (5.1), as $V_{\kappa+1} \subset M_0$ we know that M_0 agrees with V_λ on all the survival statements, i.e. we have:

$$M_0 \vDash (U_1 < U_2) \quad \text{i.e.} \quad M_0 \vDash (o(\kappa) \geq 1).$$

Now, of course we can do the usual reflection argument on the bracket determined by $j(\kappa)$ and κ in M.

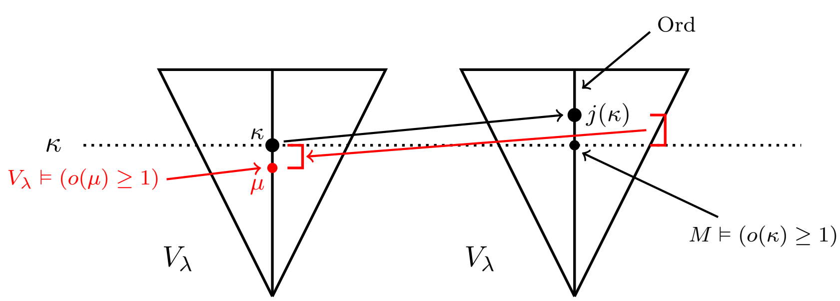

FIGURE 11. The reflection argument to find a smaller cardinal with Mitchell rank ≥ 1 when κ has Mitchell rank ≥ 2. Recall that M_0 is an inner model of V_λ, and so contains all the ordinals in V_λ.

Indeed, we know that

$$M_0 \vDash (\exists \mu)(\mu < j(\kappa) \text{ and } o(\mu) \geq 1)$$

[39]Technically, since we have not said anything regarding transitivity of this survival relation, this condition should be $U_i < U_j$ for each $i < j$.

which implies, by elementarity of the embedding,

$$V_\lambda \vDash (\exists\mu)(\mu < \kappa \text{ such that } o(\mu) \geq 1).$$

Hence, if $o(\kappa) \geq 2$, then there is a smaller cardinal which has Mitchell rank ≥ 1. Hence, if we take κ to be the smallest cardinal with $o(\kappa) \geq 2$, then we find a smaller cardinal with Mitchell rank ≥ 1, which cannot have Mitchell rank ≥ 2. Hence $o(\kappa) \geq 1 \not\Rightarrow o(\kappa) \geq 2$, and so we are done. □

5.3. **Strength.** We now move away from defining large cardinal notions by ultrafilters and use the elementary embedding itself as the object which determines the 'size' of the cardinal.

If we have any elementary embedding $j : V_\lambda \to M$ (not necessarily an ultrapower embedding), where $M \subset V_\lambda$ is a transitive inner model, and moreover j has critical point $\mathrm{crit}(j) = \kappa$, then we have seen before that this implies that $j|_{V_\kappa} = \mathrm{id}|_{V_\kappa}$, i.e. $V_\kappa \subset M$. We can then clearly strengthen this if higher levels of the von Neumann hierarchy are included in M, i.e.:

Definition 5.3. We call such an elementary embedding j *α-strong* if $V_{\kappa+\alpha} \subset M$.

We have already seen that, if κ is measurable and $j_U : V_\lambda \to M_U$ is the ultrapower embedding (where U is some κ-complete non-trivial ultrafilter on κ), then j_U is 1-strong (i.e. $V_{\kappa+1} \subset M_U$, by Claim 8), but j_U is not 2-strong (i.e. $V_{\kappa+2} \not\subset M_U$, by Theorem 4.3).

Definition 5.4. A cardinal κ is called *α-strong* if there is an α-strong elementary embedding j as above, with critical point κ.

The Fundamental Theorem of Measurable Cardinals (i.e. Theorem 4.2) along with Claim 8 gives that

$$\kappa \text{ is measurable} \iff \kappa \text{ is 1-strong}.$$

Now let us look at 2-strong cardinals. Again, by our standard reflection technique, 2-strong is (much) stronger than 1-strong (or, equivalently, measurability). Indeed, suppose that κ is 2-strong. Then we know that κ is 1-strong (and hence measurable), and so there is a κ-complete non-trivial ultrafilter U on κ (in V_λ). But as $U \in V_{\kappa+2}$, and, as κ is 2-strong, $V_{\kappa+2} \subset M$ (for M some transitive inner model of V_λ determined by some elementary embedding $j : V_\lambda \to M$), we see that M thinks that κ is measurable, i.e.

$$M \vDash (\kappa \text{ is measurable}).$$

By now we can reflect the bracket determined by $j(\kappa)$, κ, and α (for some $\alpha < \kappa$) in M back to V_λ. Indeed, we have for any $\alpha < \kappa$, that (as $j(\alpha) = \alpha$)

$$M \vDash (\exists\mu)(j(\alpha) < \mu < j(\kappa) \text{ and } \mu \text{ is measurable})$$

which implies by elementarity of j, that

$$V_\lambda \vDash (\exists\mu)(\alpha < \mu < \kappa \text{ and } \mu \text{ is measurable}).$$

Hence, if κ is 2-strong, then there are κ-many measurable cardinals $< \kappa$. Hence, is κ is the smallest 2-strong cardinal, we see that there is a measurable (i.e. 1-strong) cardinal $< \kappa$, which then cannot be 2-strong. So: 1-strong $\not\Rightarrow$ 2-strong.

It is also true that 2-strong is stronger than all of these Mitchell ranks we have seen:

Theorem 5.1. *If κ is 2-strong, then for each $n \in \mathbb{N}$ we have $o(\kappa) \geq n$.*

Proof. Not given (simply due to time constraints of the course). However, we will see a weaker version of this on Example Sheet 3; Q43 proves: *if κ is 2-strong and $o(\kappa) \geq n$, then κ is not the least cardinal with $o(\kappa)$.* $\square$

We now lift this notion of strength to a notion of being α-strong for all α:

Definition 5.5. A cardinal κ is called *strong* if it is α-strong for all ordinals α.

It is important to note that the definition of κ being strong is a separate claim for each ordinal α: we do not require that there is a witness, i.e. an elementary embedding, which works *for all* α, but just that for each α there is one. Hence, if κ is strong, then for each α there is an elementary embedding $j_\alpha : V_\lambda \to M_\alpha$ such that $V_{\kappa+\alpha} \subset M_\alpha$; we do not need to 'extend' all these embeddings into a single embedding.

Of course, one could define a stronger notion, asking for this, i.e. asking that there is a *single* embedding which works for each ordinal α. This is:

Definition 5.6. A cardinal κ is called a *Reinhardt cardinal* if there is an embedding $j : V_\lambda \to M$ with critical point κ such that j is α-strong for all ordinals α.

In particular, if κ is a Reinhardt cardinal, then we would need $V_{\kappa+\alpha} \subset M$ for all α, i.e. we would need $M = V_\lambda$. But can there even be such an elementary embedding? It turns out that there is not, and hence:

Theorem 5.2 (Kunen, 1971)**.** *There are no Reinhardt cardinals.*

This theorem is known as *Kunen's Inconsistency Theorem.* Reinhardt cardinals are (so far) the only proposed large cardinal axiom which has ever been proved to be inconsistent (in `ZFC`).

We will prove this theorem modulo a technical combinatorial lemma. Indeed, we will prove:

Theorem 5.3. *Suppose $j : V_\lambda \to M$ is an elementary embedding with $\mathrm{crit}(j) = \kappa$. Then inductively define $\kappa_0 := \kappa$, $\kappa_{i+1} := j(\kappa_i)$, and set*[40]

$$\hat{\kappa} := \bigcup_{i \in \mathbb{N}} \kappa_i.$$

There, there is $X \in V_{\hat{\kappa}+1}$ such that $X \notin M$. In particular, j is not $(\hat{\kappa}+1)$-strong.

This theorem tells us that for any elementary embedding $j : V_\lambda \to M$, there is some α for which it is not α-strong, and indeed we can even specify what α is. So, whilst there is no fixed bound on strength of a cardinal, there is a bound on the strength of a particular embedding (and hence, there are no Reinhardt cardinals, proving Theorem 5.2).

[40]This type of definition is something we are used to in finding fixed points of normal operators. In particular, $\hat{\kappa}$ is the least fixed point of j which is $> \kappa$.

Before starting the proof, let us give the statement of the combinatorial lemma which we need (but won't prove).

If X is a totally ordered set, we write $[X]^\omega$ for the set of all strictly increasing functions $\omega \to X$ (i.e. the family of subsets of X with order type ω). Then we define:

Definition 5.7. Let μ be a cardinal and $f : [\mu]^\omega \to \mu$. We say that f is *ω-Jónsson* if for all $X \subset \mu$ with $|X| = \mu$, we have

$$\{f(y) : y \in [X]^\omega\} = \mu$$

i.e. f is a function which hits every element of μ in a strong way.

The combinatorial lemma we then need is:

Lemma 5.1 (Erdös–Hajnal, 1966)**.** *Every infinite cardinal has an ω-Jónsson function.*

Proof. We do not give the proof here: the proof (which is not too complicated) can be found in Kanamori's book (*The Higher Infinite*). □

Armed now with this combinatorial lemma, we can now prove Theorem 5.3.

Proof of Theorem 5.3. Fix $j : V_\lambda \to M$ an elementary embedding with critical point $\mathrm{crit}(j) = \kappa$ with $M \subset V_\lambda$ a transitive inner model. By Erdös–Hajnal (Lemma 5.1), we may fix an ω-Jónsson function $f : [\hat{\kappa}]^\omega \to \hat{\kappa}$.

If we now look at the image of f under j, by elementarity this will be (in M) a ω-Jónsson function on $j(\hat{k})$. But $\hat{\kappa}$ was a fixed point of j, i.e. $j(\hat{\kappa}) = \hat{\kappa}$. This therefore means that, in M, $j(f)$ is also a ω-Jónsson function on $\hat{\kappa}$. Hence, we have:

$$M \vDash (j(f) \text{ is an } \omega\text{-Jónsson function on } \hat{\kappa}).$$

We need to look for a subset of $V_{\hat{\kappa}}$ which we want to not lie in M. For this, we choose: $X := \{j(\alpha) : \alpha \in \hat{\kappa}\}$. Now, because $\hat{\kappa}$ is a fixed point of j, by elementarity we get that if $\alpha \in \hat{\kappa}$, then $j(\alpha) \in j(\hat{\kappa}) = \hat{\kappa}$, and so $X \subset V_{\hat{\kappa}}$ (hence $X \in V_{\hat{\kappa}+1}$). Clearly we also have $|X| = \hat{\kappa}$.

Now, suppose for contradiction that $X \in M$. Then, by definition of $j(f)$ being ω-Jónsson in in M, we would have

$$M \vDash (\{j(f)(y) : y \in [X]^\omega\} = \hat{\kappa}). \tag{5.2}$$

Now, note that if $y \in [X]^\omega$, then y is an ω-sequence of things in X, i.e. of things of the form $j(\alpha)$, and so $y = \{j(\alpha_n) : n \in \omega\}$, where $\alpha_n \in \hat{\kappa}$. But then the function $x : \omega \to \hat{\kappa}$ sending $n \mapsto \alpha_n$ is a function with the property that $j(x) = y$.

But then $j(f)(y) = j(f)(j(x)) = j(f(x))$, where the second equality follows from elementarity of j. Thus we may write (5.2) as:

$$M \vDash (\{j(f(x)) : x \in [\hat{\kappa}]^\omega\} = \hat{\kappa}).$$

But as f is ω-Jónsson, we know that *every* element of $\hat{\kappa}$ is an value of f at some $x \in [\hat{\kappa}]^\omega$, i.e. $\{f(x) : x \in [\hat{\kappa}]^\omega\} = \hat{\kappa}$. So, this is actually just saying that:

$$M \vDash (\{j(\alpha) : \alpha \in \hat{\kappa}\} = \hat{\kappa}).$$

But this cannot be true: we know $\kappa \in \hat{\kappa}$, but κ was the critical point of j, and we know that the critical point of j will never be the j-image of anything[41], i.e. there is no $\alpha \in \hat{\kappa}$ with $\kappa = j(\alpha)$, and so $\kappa \notin \{j(\alpha) : \alpha \in \hat{\kappa}\}$, but this set was equal to $\hat{\kappa}$ which contains κ; this provides the desired contradiction. Hence $X \notin M$, which completes the proof. □

5.4. **Supercompactness.** Fix λ a cardinal and M an inner model of V_λ.

Definition 5.8. For μ an ordinal, we say that M is *closed under μ-sequences* if $M^\mu \subset M$, i.e. if every function $\mu \to M$ is an element of M.

Definition 5.9. We say that an elementary embedding $j : V_\lambda \to M$ is *μ-supercompact* if M is closed under μ-sequences.

Note that the previous notion we just saw of "α-strong" talks about the inner model M being similar to V_λ, in terms of von Neumann ranks. Supercompactness doesn't talk about von Neumann ranks, but instead talks about something which may be true everywhere in V_λ (and not just below a certain rank), and talks about the size of the sets that have to be preserved. Of course, if the size is sufficiently large it necessarily implies that all of a von Neumann rank will be preserved. Indeed, for example we can show:

Lemma 5.2. *If an elementary embedding j is 2^κ-supercompact, where crit$(j) = \kappa$, then j is 2-strong.*

Proof. If $X \in V_{\kappa+2}$, then $X \subset V_{\kappa+1}$, and so $|X| \leq |V_{\kappa+1}| = 2^{|V_\kappa|} = 2^\kappa$, where in this last inequality we have used the fact that $|V_\kappa| = \kappa$, which is true as κ is strongly inaccessible.

Therefore, we can think of every element of $V_{\kappa+2}$ as a sequence of length 2^κ, i.e. since $V_{\kappa+1} \subset M$[42], we have $X \subset M$, and so X is given by a function $2^\kappa \to M$. Hence, as j is 2^κ-supercompact, we have that $X \in M$. This therefore shows that $V_{\kappa+2} \subset M$, i.e. κ is 2-strong. □

So, a sufficient amount of supercompactness will give a certain amount of strength.

Our next theorem tells us that in the specific case of the ultrapower embedding j of a measurable cardinal (as constructed in Section 4), this M is closed under κ-sequences, and hence this j is κ-supercompact:

Theorem 5.4. *If κ is measurable and M is the ultrapower (relative to a given κ-complete non-trivial ultrafilter on κ), then $M^\kappa \subset M$.*

[41] Indeed, we know that $j|_{V_\kappa} = \mathrm{id}|_{V_\kappa}$ and $j(\kappa) > \kappa$, so κ cannot be in the image of j.
[42] The proof of Claim 8 still works here.

Proof. If $x \in M^\kappa$, then x is a function $x : \kappa \to M$, and so we can write x as a sequence of elements of M, i.e. $x = ((f_\alpha)_U : \alpha < \kappa)$, where $f_\alpha : \kappa \to V_\lambda$ is a representative of α^{th}-term in x.

How could we describe this in M? We know that there is something which describes κ in M; indeed, we know $\kappa \in M$, and so let $h : \kappa \to V_\lambda$ be such that $(h)_U = \kappa$ (in fact, as before we know that $h : \kappa \to \kappa$).

Now define a function g with the following properties: for each $\xi \in \kappa$, $g(\xi)$ is the function g_ξ such that:

- $\mathrm{dom}(g_\xi) = h(\xi)$ (note that $h(\xi) \in \kappa$);
- for each $\alpha \in \mathrm{dom}(g_\xi)$, we set: $g_\xi(\alpha) := f_\alpha(\xi)$.

Note that for each $\xi \in \kappa$, we have $g_\xi : h(\xi) \to V_\lambda$. Now, as for each $\xi < \kappa$, $g(\xi)$ is a function, Łoś's theorem gives that $(g)_U$ is also a function (which moreover is an element of M). But for every $\xi \in \kappa$ we have $\mathrm{dom}(g_\xi) = h(\xi)$, and again Łoś's theorem gives that $\mathrm{dom}((g)_U) = (h)_U$, which by definition $(h)_U = \kappa$, i.e. $\mathrm{dom}((g)_U) = \kappa$. Moreover, if $\alpha < \kappa$, Łoś's theorem gives that $(g)_U(\alpha) = (f_\alpha)$. All of this together means that we have exactly $(g)_U = x$, and hence as $(g)_U \in M$, we have $x \in M$. This completes the proof. □

Let us now define what supercompactness means for cardinals:

Definition 5.10. For μ an ordinal, a cardinal κ is called *μ-supercompact* if there is an elementary embedding $j : V_\lambda \to M$ such that j is μ-supercompact.

Definition 5.11. A cardinal κ is called *supercompact* if it is μ-supercompact for all ordinals μ.

Theorem 5.4 tells us that:

$$\kappa \text{ is measurable} \iff \kappa \text{ is } \kappa\text{-supercompact}.$$

However, the ultrapower embedding of a measurable cardinal is *not* κ^+-supercompact (see Example Sheet 3, Q44).

Kunen's Inconsistency Theorem (Theorem 5.2) tells us that for every <u>fixed</u> elementary embedding, there is an upper bound (depending on the embedding) for supercompactness. So, we actually find from Kunen's Inconsistency that there is no embedding $j : V_\lambda \to M$ with $\mathrm{crit}(j) = \kappa$ that is $\hat{\kappa}$-supercompact (where $\hat{\kappa}$ is as in Theorem 5.3). This again gives that to be supercompact there must be different witnesses for the different μ-supercompactness statements.

6. Tying Up Loose Ends

In this last section, we will look comment on three different topics, namely:

(1) The relationship between large cardinals and GCH;

(2) Witness objects (namely, to try and conceptualise what is really happening in reflection proof). Here, we will also discuss why the Fundamental Theorem of Measurable Cardinals is such an important tool (this is because it provides witness objects); we will also comment on various witness objects for other stronger large cardinal notions.

(3) The annoying assumption that we had to always make in Section 4 concerning the existence of an inaccessible cardinal above the measurable cardinal.

6.1. **Large Cardinals and GCH.** The following is a classical result from the early-modern days of large cardinals:

> **Theorem 6.1** (Dana Scott). *If κ is measurable and GCH holds below κ (i.e. for all $\mu < \kappa$, $2^\mu = \mu^+$), then $2^\kappa = \kappa^+$.*

Thus, this theorem tells us that measurable cardinals can never be the first counterexample to GCH.

Proof. This follows directly from the ultrapower analysis we have already done.

If GCH holds below κ, then by elementarity of the ultrapower embedding, this can be moved up to $j(\kappa)$ in M, i.e.

$$M \vDash (\texttt{GCH}\text{ holds below } j(\kappa)).$$

So, in particular, as $\kappa < j(\kappa)$, we have $M \vDash (2^\kappa = \kappa^+)$.

Now, from our analysis, we know that $V_{\kappa+1} \subset M$, and therefore $\mathcal{P}(\kappa) \subset M$, and so therefore whatever M thinks is 2^κ, V_λ also thinks is 2^κ. Equally, as κ^+ is the supremum of all well-ordering on κ, which are all subsets of κ, κ^+ in M is the same as that in V_λ. Hence (writing α^M to represent the value of a cardinal in M):

$$\kappa^+ = (\kappa^+)^M = (2^\kappa)^M = 2^\kappa.$$

□

How does this continue? Let us see how we can get even more out of the stronger large cardinal axioms, namely for supercompact cardinals.

> **Proposition 6.1.** *If κ is γ-supercompact and μ is such that $\mu \leq \gamma$, and if GCH holds below κ, then $2^\mu = \mu^+$.*

Proof. This is almost the same proof as Theorem 6.1, except we need to ensure that we can use $j(\kappa)$ as an upper bound (as we don't know in advance what the relationship between $j(\kappa)$ and γ is, and we need $\gamma < j(\kappa)$ for that argument).

The argument that we gave in Kunen's Inconsistency Theorem tells us that, without loss of generality, we can assume that there is an elementary embedding where $\gamma < j(\kappa)$: indeed, we showed there that if we have an elementary embedding j with $\text{crit}(j) = \kappa$, then j cannot be $\hat{\kappa}$-supercompact. But here we know that it is γ-supercompact by assumption, and so we need $\gamma < \hat{\kappa}$. By definition of $\hat{\kappa}$, this therefore tells us that we must have $\gamma < j^{(n)}(\kappa)$ for some $n \in \omega$.

But then if we set $\tilde{j} := j^{(n)}$, this is a witness for the elementary embedding being γ-supercompact (i.e., it also is), and moreover $\gamma < \tilde{j}(\kappa)$. Working with $\tilde{j}$ instead, we may therefore without loss of generality assume that $\gamma < j(\kappa)$.

But now we may give the same argument as in Theorem 6.1. Indeed, because `GCH` holds below κ, elementarity of j gives that $M \vDash$ (`GCH` holds below $j(\kappa)$), and thus as $\mu \leq \gamma < j(\kappa)$, we have $M \vDash (2^\mu = \mu^+)$.

But the γ-supercompactness of j (and $\gamma < j(\kappa)$) gives that $\mathcal{P}(\mu) \subset M$, and hence $(2^\mu)^M = 2^\mu$ and $(\mu^+)^M = \mu$, and so $2^\mu = (2^\mu)^M = (\mu^+)^M = \mu^+$, as desired. □

We remark that these results were proved in the 1960's as there was a movement called *Gödel's program*: Gödel had suggested that large cardinals might actually be a way to resolve the Continuum Problem, and there was a hope that if you add large cardinals you might actually be able to either prove or refute the Continuum Problem. So questions of the above types were important in figuring out the relationship between large cardinals and the validity of `GCH`.

6.2. **Witness Objects.** Recall that the fundamental theorem of measurable cardinals (Theorem 4.2) was an equivalence between two *existential* statements (i.e. a certain ultrafilter U on κ, and an elementary embedding with critical point κ). However, the objects which are claimed to exist are rather different: the ultrafilter U is an element of $V_{\kappa+2}$, whilst the elementary embedding j is an element of $V_{\lambda+1}$.

This means that the power of the fundamental theorem is that you have this very powerful statement of model theory or set theory (elementary embeddings), which is powerful as it is a statement about the whole universe V_λ, and you show that it is equivalent to something which is bounded well below V_λ, namely something in $V_{\kappa+2}$. This is remarkable that you can claim something about the whole universe just from the existence of an element much lower done.

Thus: the power of the fundamental theory is that the existence of such an elementary embedding is actually witnessed by an object in $V_{\kappa+2}$.

Reflection arguments are based on this. For example, if κ is 2-strong, then we saw that κ is the κ^{th}-measurable. This was one of our reflection arguments, and the reason it works is that measurability is witnessed by something that lives in $V_{\kappa+2}$, and therefore the 2-strong embedding preserves that.

Let us briefly discuss, without proofs, the appropriate witness objects for the strong large cardinal notions we have seen, namely supercompactness, strength, and strong compactness.

6.2.1. *Witness Objects for Supercompactness.* Theorem 22.7 in Kanamori's book, *The Higher Infinite*, says:

> **Theorem 6.2.** *If $\kappa \leq \gamma$, then κ is γ-supercompact if and only if there is a normal[43] ultrafilter over $\mathcal{P}_\kappa(\gamma) := \{X \in \mathcal{P}(\gamma) : |X| \leq \kappa\}$.*

In particular, where does this witness object live, i.e. where does a (normal) ultrafilter on $\mathcal{P}_\kappa(\gamma)$ live? Well, it is readily seen to live in $V_{\gamma+2}$, and so the above theorem is like the fundamental theorem of measurable cardinals as it reduces a statement about elementarity (which is really a statement about subsets of V_λ) to a statement talking about objects in $V_{\gamma+2}$.

We also note that Exercise 22.8 in Kanamori's book says:

(a) If κ is supercompact and $\lambda > \kappa$ is inaccessible, then $V_\lambda \vDash$ "κ is supercompact".

(b) If $\texttt{ZFC} + (\exists\kappa)(\kappa$ is supercompact) is consistent, then so is

$$\texttt{ZFC} + (\exists\kappa)(\kappa \text{ is supercompact, and there is } \underline{\text{no}} \text{ inaccessible } \lambda > \kappa).$$

We mention this to highlight that this business about witness objects is also behind our intuition that adding an inaccessible cardinal above actually gives us a stronger theory. Recall that we started the course by showing that some theories are stronger than others because can "cut off" the universe at any inaccessible and then you get something that does not contain that inaccessible anymore. Hence, we can't always prove that there is an inaccessible above a certain cardinal; hence our assumption that there is one is a stronger theory.

Now, it is not clear that notions such as supercompactness even have this property, because full supercompactness said for all ordinals γ, there is a γ-supercompact embedding. So, it is not at all clear that if we have a supercompact and a inaccessible above it, that this gives us a stronger theory. But now, the moment you have witness objects it is clear: indeed, the above exercises show that the assumption:

$$(\exists\kappa)(\exists\lambda)((\kappa < \lambda) \wedge (\kappa \text{ has property } \Phi) \wedge (\lambda \text{ is (strongly) inaccessible}))$$

where Φ is some cardinal property, is strictly stronger than just $\Phi\texttt{C}$, i.e. $(\exists\kappa)(\kappa$ has property $\Phi)$.

6.2.2. *Witness Objects for Strongly Compact Cardinals.* Here we need a slightly different notion of filter known as a *fine* filter (which we also won't define). Then:

Definition 6.1. A cardinal κ is called *γ-compact* if there is a fine ultrafilter on $\mathcal{P}_\kappa(\gamma)$.

Then in Kanamori, Theorem 22.17 we have:

Theorem 6.3. *Let $\kappa \leq \gamma$. Then, the following are equivalent:*

(a) *κ is γ-compact;*

(b) *there is an elementary embedding $j : V \to M$ with crit$(j) = \kappa$ such that for any $X \subset M$ with $|X| \leq \gamma$, there is a $Y \in M$ with $X \subset Y$ and $M \vDash (|Y| < j(\kappa))$;*

(c) *For any set S, every κ-complete filter over S which is generated by at most $|\gamma|$ sets can be extended to a κ-complete ultrafilter over S.*

Note that (b) is weaker than supercompactness, and (c) is a type of extendibility statement.

To get that this really is strong compactness as we know it, we would need to link statement (c) regarding "extending filters to ultrafilters" to that of strong compactness. If you recall the

[43] We won't define what "normal" means in this context.

Keisler–Tarski Theorem (Theorem 2.4), which said that if κ is strong compact then every κ-complete filter over any set can be extended to a κ-complete ultrafilter, to get the equivalence with the above theorem we would need the converse of Keisler–Tarski. The proof of this is in Kanamori's book (Page 37, Proposition 4.1).

6.2.3. *Witness Objects for Strong Cardinals.* The witness objects for strong cardinals are somewhat more complicated than the others – given a longer course, they could easier be introduced, but they would be the focus of an entire topic in itself. However, we will mention their name and a rough definition as they are some of the most important objects for modern research into large cardinal properties. The notion is that of an *extender*.

Definition 6.2. A (κ, γ)-*extender* is a family of ultrafilters $\{E_a : a \subset \gamma \text{ is finite}\}$ with certain cohesion properties.

These extenders can be used as witness objects for strong embeddings – see Section 26 in Kanamori's book.

6.3. **The Annoying Inaccessible Above.** Recall that the fundamental theorem of measurable cardinals (Theorem 4.2) roughly says that the following are equivalent:

(i) there exists an elementary embedding $j : V_\lambda \to M$ with $\mathrm{crit}(j) = \kappa$;

(ii) there exists U a κ-complete non-trivial ultrafilter on κ.

What we would like is to have (i) not assuming that there is an inaccessible cardinal λ ($> \kappa$). In order to remove this inaccessible, we would like to be able to express, in a model of set theory, that there is such an elementary embedding to an inner model (of the original model).

This is why we highlighted in Section 6.2 that (i) is really a statement about $V_{\lambda+1}$, as this is where the elementary embedding j lies. This means that, if V_λ is our model of set theory in which we are working, there is no chance we can actually talk about j itself. That is what makes this difficult, and this is why we ignored this initially by working in a bigger model where in the bigger model we can talk about this j as it is just an object in the bigger model, which made everything easier. For example, we could take the ultrapower of it as it was a set, and we didn't have to worry about thinking what this actually means.

Now, of course this j we are talking about, in a given model of set theory, would not be a set (in general). It would be a proper class, and of course set theorists don't really like to talk about them a lot, as that is where set theory comes to its own limits.

So, if we are working in a fixed model $V \vDash \mathtt{ZFC}$, then both the elementary embedding j and the inner model M would be proper classes, and therefore we cannot quantify over them and can't write things like "$(\exists j)(\cdots)$".

Let us say a few words about classes. There are various ways that you can try to formalise classes: if you have a model of set theory and a "meta-theory" and you look from the outside at this model, then its classes are really just the "subsets" of that model, i.e. if our model $V \vDash \mathtt{ZFC}$ is a "set" in our meta-set theory, then its classes are just $\mathcal{P}(V)$ (the "meta-power set"), and so in particular there are many more classes than sets (as there are many sets as V has elements, and there are $2^{|V|}$ classes).

We know that a lot of these classes are relatively unproblematic, and we deal with them all the time (for example, the class of all ordinals, the class of all sets). So, many classes that matter in practice are definable (and often absolutely definable): indeed, if ϕ is a formula with $n+1$ free variables, and $a \in V^n$, then we can apply "meta-separation" to get a class

$$C_{\phi,a} := \{x \in V : V \vDash \phi(x,a)\}.$$

Of course, there are far fewer definable classes than classes itself (as there are only $|V|$ definable classes over V, but $2^{|V|}$ classes over V).

So can we now rephrase our fundamental theorem, as a meta-theorem about models of set theory, by talking about definable classes? Well, it would be:

(Reformulation of the fundamental theorem of measurable cardinals.) *The following are equivalent:*

(i) there is ϕ and a such that $C_{\phi,a}$ is an elementary embedding with critical point κ[44];

(ii) there exists U a κ-complete non-trivial ultrafilter over κ[45].

> **Remark:** An alternative approach to doing this as a meta-theorem is to not do set theory, but to actually do *class theory* (where you would have a two-type language, for sets and classes). Then, you could formulate this in a first-order manner in some standard class theory, such as NBG (von Neumann–Bernays–Gödel).

Let us now look at the proof of this new version of the fundamental theorem, and in particular see if our original proof still works.

(i)$\Rightarrow$(ii): This is not really a problem, because the proof involved a concrete definition of an ultrafilter, and so this works exactly as before, replacing all references to j by appropriate terms involving ϕ and a.

(ii)$\Rightarrow$(i): Given an ultrafilter U, we want to give a formula, using U as a parameter, which defines the embedding and the inner model, i.e. a formula Φ such that $x \in M \Leftrightarrow \Phi(x)$ (once we have the inner model, the embedding is straightforward to define). What would Φ need to be? Well, if we look at our original proof of Theorem 4.2, we would need:

$$\Phi(x) \quad :\Longleftrightarrow \quad x \text{ is the image of the Mostowski collapse of some } [f]_{\sim_U}, \text{ where } f : \kappa \to V$$

(note that a function from $f : \kappa \to V$ is fine, as V is a model of replacement and so $f \in V$).

But the real problem here is that the equivalence class is a proper class and really is not a set: so $[f]_{\sim_U}$ is not an element of V[46].

---

[44] To stress, the "there is" here is in the meta-language. The condition that there is an "elementary embedding" would be infinitely many formulas in the meta-theory, as we need each formula to be preserved by the embedding.

[45] Note that the set-theoretic part of this equivalence does not change.

[46] This is because you can change the values of f in a single value without changing the equivalence class, and so you can just put in at the value 0 an arbitrary set, and so the equivalence class will be unbounded in the ranks of V, i.e. if f is any function and $x \in V$ is any set, then

$$f_x : \alpha \mapsto \begin{cases} x & \text{if } \alpha = 0; \\ f(\alpha) & \text{if } \alpha \neq 0 \end{cases}$$

has $f_x \sim_U f$, but the ranks of the f_x are unbounded in V. Hence, $[f]_{\sim_U}$ really is a proper class for each f.

So therefore you can't really say "there is an equivalence class" for the same reason as before, as this would be quantifying over a class. The solution to this is to use the axiom of foundation in V via something known as *Scott's trick*:

Consider the class $[f]_U \subset V$. Find α an ordinal minimal such that $[f]_U \cap V_\alpha \neq \emptyset$. Now, this intersection $[f]_U \cap V_\alpha$ *is* a set in V, and so we write $\mathrm{scott}(f) := [f]_U \cap V_\alpha$ for this set.

Clearly, if $f \not\sim_U g$, then $\mathrm{scott}(f) \cap \mathrm{scott}(g) = \emptyset$. This little trick allows us to now write down the formula Φ we want: set

$$\Phi(x) \quad :\Longleftrightarrow \quad x \text{ is the Mostowski collapse of some } \mathrm{scott}(f) \text{ for some } f : \kappa \to V$$

i.e. replace $[f]_{\sim_U}$ in our previous definition of Φ by $\mathrm{scott}(f)$. This then allows us to define the ultrapower of the universe, and therefore we can get the fundamental theorem back, thus dealing with this issue regarding "there exists an inaccessible above".

End of Lecture Course

Example Sheet 1

(1) If α is an ordinal, the *ordinal topology* on α is the topology generated by the basic open sets $L_\beta := \{\gamma \in \alpha : \gamma < \beta\}$ and $R_\beta := \{\gamma \in \alpha : \gamma > \beta\}$. Check that every successor ordinal $\beta + 1 \in \alpha$ is an isolated point in this topology and determine the neighbourhoods of a limit ordinal $\lambda \in \alpha$.

(2) We call the axiom system that contains all axioms of `ZFC` except for the Axiom of Infinity `FST` (for "Finite Set Theory"). Consider the property $I(\alpha)$ defined by "α is a limit ordinal and $\alpha \neq 0$". Show that the property I is a *large cardinal property* for FST in the following sense: *If* `FST` *is consistent, then* `FST` *does not prove the existence of a cardinal with property I.*

(3) Let λ and μ be limit ordinals and $f : \mu \to \lambda$ a function. The function f is called *cofinal* in λ if $\operatorname{ran}(f)$ is a confinal subset of λ. Show that

$$\begin{aligned}\operatorname{cf}(\lambda) &= \min\{\mu : \text{ there is a cofinal function with domain } \mu\}\\ &= \min\{\mu : \text{ there is a strictly increasing cofinal function with domain } \mu\}.\end{aligned}$$

Conclude that $\operatorname{cf}(\operatorname{cf}(\lambda)) = \operatorname{cf}(\lambda)$.

(4) Let κ be regular, η any ordinal, and $f : \kappa \to \eta$ a strictly increasing function. Define $\lambda := \bigcup \operatorname{ran}(f)$. Show that $\operatorname{cf}(\lambda) = \kappa$. Conclude that $\operatorname{cf}(\aleph_\lambda) = \operatorname{cf}(\lambda)$.

(5) We said that a cardinal κ satisfies *second order replacement* if for all $G : V_\kappa \to V_\kappa$ and $x \in V_\kappa$, the set $G[x] := \{G(y) : y \in x\} \in V_\kappa$. In the proof of Theorem 1.3, we showed that if κ is inaccessible, then it satisfies second order replacement. Show the converse. (This is known as *Shepherdson's Theorem.*)

(6) Let κ be a regular cardinal. If x is any set, we write $\operatorname{tcl}(x)$ for the transitive closure of x. Define $H_k := \{x : |\operatorname{tcl}(x)| < \kappa\}$. Show that $V_\kappa = H_\kappa$ if and only if κ is inaccessible.

(7) Suppose that $(M, \in)$ and $(N, \in)$ are models of `ZFC` with $M \subset N$ and M is transitive in N. Show that the notions of "function", "injection", "surjection", "bijection", and "cofinal" are absolute between M and N.

(8) Let κ be inaccessible and $\lambda < \kappa$. Show that λ is inaccessible if and only if $V_\kappa \models$ "λ is inaccessible".

(9) Show that every worldly cardinal is a limit cardinal.[47]

(10) Prove the *Tarski–Vaught Test* for being an elementary substructure (i.e. Proposition 1.3).

(11) Prove *Tarski's Chain Lemma* (see before Theorem 1.6).

(12) Let β be any ordinal and $R \subset V_\beta$. An ordinal $\alpha < \beta$ is called an *R-Lévy ordinal for β* if $(V_\alpha, \in, R \cap V_\alpha)$ is an elementary substructure of $(V_\beta, \in, R)$. Show that no α can be an R-Lévy ordinal for all $R \subset V_\beta$.

(13) Show the following theorem due to Lévy: an ordinal κ is an inaccessible cardinal if and only if for each $R \subset V_\kappa$ there is an R-Lévy ordinal for κ.

[47] *Hint:* Use the fact that the proof of Hartog's Lemma implies that there is a surjection from the power set of κ onto κ^+.

Example Sheet 2

(14) Modify the proof that ZFC (if consistent) does not prove IC to a proof of "if ZFC +GCH is consistent, then ZFC does not prove that there are weakly inaccessible cardinals". Argue that this gives rise to a proof of the unprovability of the existence of weakly inaccessibles that does not need all of Gödel's 1938 theorem.

(15) Let 2IC be the statement "there are $\lambda < \kappa$ such that both λ and κ are inaccessible". Show that if $\mathtt{ZFC} + \mathtt{IC}$ is consistent, then IC does not imply 2IC.

(16) Show that there is a Π_1 formula ϕ such that $\mathtt{ZFC} \vdash \phi(x)$ if and only if x is a strong limit cardinal.

(17) Remind yourself of Mostowski's Collapsing Theorem (see the notes on *Logic and Set Theory*). Let κ be inaccessible. In the current notes, we constructed a countable, non-transitive $M \subset V_\kappa$ such that $M \preceq V_\kappa$. Use Mostowski's Collapsing Theorem to show that there is a transitive set $M^* \in V_\kappa$ such that $(M^*, \in)$ is isomorphic to $(M, \in)$. In particular, $M^* \subset V_\kappa$ is a transitive submodel of ZFC.

(18) Using the model M^* from (17), explain why Π_1 formulas are not in general absolute between transitive models of ZFC.[48]

(19) Show that the smallest Ulam cardinal is a measurable cardinal.

(20) Suppose $\mu : \kappa \to 2$ and $U \subset \mathcal{P}(\kappa)$; define $\mu_U(A) := 1$ if $A \in U$ and $U_\mu := \{A : \mu(A) = 1\}$. Show that if U is a κ-complete non-trivial ultrafilter on κ, then μ_U is a κ-additive non-trivial measure on κ, and if μ is a κ-additive non-trivial measure on κ, then U_μ is a κ-complete non-trivial ultrafilter on κ.

(21) Let κ be regular. Show that $\{X : |\kappa \setminus X| < \kappa\}$ is a κ-complete filter that is not an ultrafilter.

(22) Using the Axiom of Choice, show that every filter can be extended to an ultrafilter (preserving non-triviality).

(23) A model $(M, E) \vDash \mathtt{ZFC}$ is called an *ω-model* if its natural numbers are standard, i.e. if there is an isomorphism between $(\{x \in M : M \vDash \text{“}x \text{ is a natural number”}\}, E)$ and $(\omega, \in)$. Let M be an ω-model; without loss of generality, we can assume that $\omega \subset M$. We encode formulas of first-order logic by natural numbers, writing $\ulcorner\phi\urcorner$ for the number coding ϕ. Let Φ be a set of first-order sentences such that Φ exists in M, i.e., there is some $x \in M$ such that $\phi \in \Phi$ if and only if $M \vDash \ulcorner\phi\urcorner \in x$. Show that Φ is consistent if and only if $M \vDash$ "Φ is consistent". Deduce that if $\mathtt{ZFC} + \mathrm{Cons}(\mathtt{ZFC})$ is consistent, it cannot show the existence of an ω-model.

(24) Find an $\mathcal{L}_{\omega_1,\omega}$ formula that characterises the ω-models of ZFC.

(25) Give a concrete uncountable collection of $\mathcal{L}_{\omega_1,\omega}$ sentences that is countably satisfiable, but not satisfiable.

(26) If κ is a strongly compact cardinal, the Keisler–Tarski theorem makes a statement about κ-complete filters on arbitrary sets X. What does the proof show if κ is only assumed to be weakly compact? Why is that useless?[49]

[48] *Hint:* What is $\mathrm{Ord} \cap M^*$? If $\kappa \in M^*$ is such that $M^* \vDash$ "κ is a cardinal", can κ be a real cardinal?

[49] *Hint:* If $\lambda < \kappa$, which filters on λ can be κ-complete?

(27) In a reflection argument, we used Keisler's Theorem on the Extension Property to show that below each weakly compact cardinal is an inaccessible by reflecting the property "κ is inaccessible". Clearly, it cannot be possible to reflect the property "κ is weakly compact". Explain where the argument breaks down if you try to prove this.

(28) Let $\infty\mathtt{IC}$ be the statement "for all ordinals α, there is $\kappa > \alpha$ such that κ is inaccessible". Show that if κ is weakly compact, then $V_\kappa \vDash \infty\mathtt{IC}$.

(29) Suppose that κ is a measurable cardinal and U is a κ-complete ultrafilter on κ, and $\pi : V_\kappa \to \mathrm{Ult}(V_\kappa, U)$ is the ultrapower embedding, i.e. $\pi(x) := (c_x)_U$. By Łoś's Theorem, π is an elementary embedding. Show that $\{\pi(x) : x \in V_\kappa\}$ is isomorphic to V_κ and transitive in $\mathrm{Ult}(V_\kappa, U)$, i.e., if $z \in \pi(x)$, then there is $y \in V_\kappa$ such that $z = \pi(y)$.

Conclude that the order type of the ordinals in $\mathrm{Ult}(V_\kappa, U)$ is not equal to κ and that therefore $\mathrm{Ult}(V_\kappa, U)$ is not isomorphic to V_κ.

Example Sheet 3

(30) Let κ be inaccessible and L be any $\mathcal{L}_{\kappa,\kappa}$ language with M an L-structure. Write L^α for the set of L-formulas whose free variables are contained in $\{v_\xi : \xi < \alpha\}$. If $X \subset M$, we say that X is an *L-elementary substructure* (in symbols: $X \preceq_L M$) if for all $\phi \in L^\alpha$ and all $x \in X^\alpha$, we have that

$$X[x/v] \vDash \phi \qquad \Longleftrightarrow \qquad M[x/v] \vDash \phi.$$

Prove the following statement (the *Tarski–Vaught Test for $\mathcal{L}_{\kappa,\kappa}$ languages*): a subset X is an L-elementary substructure if and only if it is an L-substructure and for all $\phi \in L^{\alpha+\beta}$ (with $v := \{v_\xi : \xi < \alpha\}$ and $w := \{v_{\alpha+\eta} : \eta < \beta\}$) and all $x \in X^\alpha$, if $M[x/v] \vDash (\exists^\beta w)\phi$, then there is some $y \in X^\beta$ such that $M[x/v, y/w] \vDash \phi$.

(Why do we require the inaccessibility of κ?)

(31) Let κ be inaccessible, L any $\mathcal{L}_{\kappa,\kappa}$ language, M an L-structure, and $X \subset M$. If $\phi \in L^{\alpha+\beta}$ (with $v := \{v_\xi : \xi < \alpha\}$ and $w := \{v_{\alpha+\eta} : \eta < \beta\}$) and $x \in M^\alpha$ such that $M[x/v] \vDash (\exists^\beta w)\phi$, then there is some $y \in M^\beta$ such that $M[x/v, y/w] \vDash \phi$. Use the Axiom of Choice to assign such a witness $w(\phi, x)$. Let $H(X, \alpha) := X \cup \bigcup\{\mathrm{ran}(w(\phi, x)) : \phi \in L^{\alpha+\beta}, x \in X^\alpha\}$. Define by recursion $H_0(X) := X$, $H_{\alpha+1}(X) := H(H_\alpha(X, \alpha))$, and $H_\lambda(X) := \bigcup_{\alpha<\lambda} H_\alpha(X)$ (for non-zero limit ordinals λ) and show that $H_\kappa(X) \preceq_L M$ is an L-elementary substructure of cardinality $\leq \kappa$.

(32) Show that the consistency strength hierarchy has the following properties:

(a) $(0 = 1)$ is maximal with respect to $\leq_{\text{Cons}}$;

(b) If $\mathtt{A}$ is not maximal, then there is $\mathtt{B}$ such that $\mathtt{A} <_{\text{Cons}} \mathtt{B}$ and $\mathtt{B}$ is not maximal;

(c) for all $\mathtt{A}$ and $\mathtt{B}$, if $\mathtt{A} \leq_{\text{Cons}} \mathtt{B}$, then $\mathtt{A} \vee \mathtt{B} \equiv_{\text{Cons}} \mathtt{A}$.

(33) Let Φ be a cardinal property (i.e. $\Phi(\kappa)$ implies that κ is a cardinal). Let us say that Φ is *non-trivial* if $\Phi(\kappa)$ implies that κ is inaccessible. Show that there is a non-trivial Φ such that $\Phi\mathtt{C} \equiv_{\text{Cons}} \mathtt{IC}$ and $\mathtt{WC} <_1 \Phi\mathtt{C}$. Use this to argue that the following statement is in general false: if $\mathtt{A} \leq_{\text{Cons}} \mathtt{B}$, then $\mathtt{A} \vee \mathtt{B} \equiv_{\text{Cons}} \mathtt{B}$.

(34) Let $\mathtt{A}$ be the statement "if there is a weakly compact cardinal κ, then there is an inaccessible $\lambda > \kappa$". Show that the consistency strength of $\mathtt{ZFC} + \mathtt{A}$ is equal to that of $\mathtt{ZFC}$, but that under some consistency assumptions, $\mathtt{ZFC} <_0 \mathtt{ZFC} + \mathtt{A}$. What are the required consistency assumptions for the latter claim?

(35) Suppose that there are unboundedly many inaccessible cardinals. Let ι_α be the α^{th} inaccessible cardinal. Show that it is not possible to prove[50] that the operation $\alpha \mapsto \iota_\alpha$ has a fixed point, i.e., some $\kappa = \iota_\kappa$. This must mean that the operation is in general not a normal ordinal operation. What is the reason?

(36) Show that if U is an ultrafilter, then U is free if and only if U is non-trivial.

(37) Let λ be inaccessible and $M \subset V_\lambda$ a transitive set. Suppose $j : V_\lambda \to M$ is an elementary embedding. Show that if $j \neq \mathrm{id}$, then there is an ordinal α such that $j(\alpha) > \alpha$.

[50] In $\mathtt{ZFC}$ + "there are unboundedly many inaccessible cardinals".

(38) We assume that $\kappa < \lambda$ are measurable and inaccessible, respectively, and that $j : V_\lambda \to M$ is the ultrapower embedding. In the notes, we showed that $\kappa \leq (\mathrm{id}) < j(\kappa)$. Give concrete functions $f : \kappa \to \kappa$ such that $(f) = (\mathrm{id})+1$, $(f) = (\mathrm{id})+\omega_1$, and $(f) = (\mathrm{id})\cdot 2$. Fix $\xi < \kappa$ and consider the function $f(\alpha) := \xi$ if α is even and $f(\alpha) := \alpha$ if α is odd[51]. What can we say about the relation between (id) and (f)?

(39) Let κ be measurable. Show that there is some ultrafilter U on κ such that in the ultrapower M_U, we have that $\kappa = (\mathrm{id})_U$, where $\mathrm{id} : \kappa \to \kappa$ is the map $\alpha \mapsto \alpha$.

(40) Let κ be a cardinal. We say κ is (defining inductively):

- 0-*inaccessible* if κ is inaccessible;
- $(\alpha + 1)$-*inaccessible* if κ is α-inaccessible and $\{\mu < \kappa : \mu \text{ is } \alpha\text{-inaccessible}\}$ is unbounded in κ; and
- λ-*inaccessible* (for λ a non-zero limit ordinal) if κ is α-inaccessible for all $\alpha < \lambda$.

Show that every measurable cardinal κ is κ-inaccessible.

(41) Let λ be inaccessible. Suppose that $M \subset V_\lambda$ is an inner model of `ZFC` closed under κ-sequences (i.e., $M^\kappa \subset M$) with $V_{\kappa+1} \subset M$, L is a language with at most κ many non-logical symbols, and that N is an L-structure with $|N| \leq \kappa$. Show that there is some $\overline{N} \in M$ such that N and $\overline{N}$ are isomorphic. Use this and Q(31) to finish the proof of Theorem 4.4 that a measurable cardinal remains weakly compact in the ultrapower.

(42) We showed in Section 5.1 that if κ is surviving, there are functions f and g such that

$$M_U \vDash \text{“}(g)_U \text{ is an } (f)_U\text{-complete ultrafilter on } (f)_U\text{”}.$$

Use this to give an alternative proof of the fact that a surviving cardinal κ must be the κ^{th} measurable cardinal.

(43) Show that if κ is 2-strong and satisfies $o(\kappa) \geq n$, then there are unboundedly many cardinals $\lambda < \kappa$ such that $o(\lambda) \geq n$.

(44) Let κ be measurable and M the ultrapower built from a κ-complete ultrafilter on κ. Show that M is not closed under κ^+-sequences by producing a function $f : \kappa^+ \to M$ that is not an element of M.

[51] As usual, an ordinal α is *even* if it is of the form $\lambda + 2n$ where λ is a limit ordinal and n is a natural number.