The Theft and the Honest Toil: Applications of Large Cardinal Axioms to the Theory of Measurable Selection

Dmitriy Kunisky
Class of 2014

May 5, 2014

Advisor: John P. Burgess (Dept. of Philosophy)

Second Reader: Zeev Dvir (Dept. of Mathematics)

Submitted to the Department of Mathematics in partial fulfillment of requirements for the A.B. degree at Princeton University.

Contents

1 Introduction

The method of 'postulating' what we want has many advantages; they are the same as the advantages of theft over honest toil.
— Bertrand Russell, *Introduction to Mathematical Philosophy*

1.1 The Problem of Selection

The problem of selection, in its most general formulation, is that of making many simultaneous choices from an indexed collection of sets of options: for sets X, Y and a subset $A \subseteq X \times Y$ (in this kind of setting we will often appeal to visual intuition and call A a *planar set*), we write $A(x)$ for the section of A over x,

$$A(x) = \{y \in Y \mid (x, y) \in A\}.$$

Thus, we view A as a parametrized sequence of subsets of Y (we will later introduce multifunctions, which make this viewpoint more explicit). If each $A(x)$ is non-empty, then the Axiom of Choice guarantees that there is a function $f : X \to Y$ such that for all $x \in X$, $f(x) \in A(x)$. That is, f "selects" one value from $A(x)$ for each x; we call any such function a *selection* of A, and we call its graph $\mathrm{Gr}(f)$ (as a planar set) a *uniformization* of A. We will also sometimes refer to A as the *choice set*. In prac-

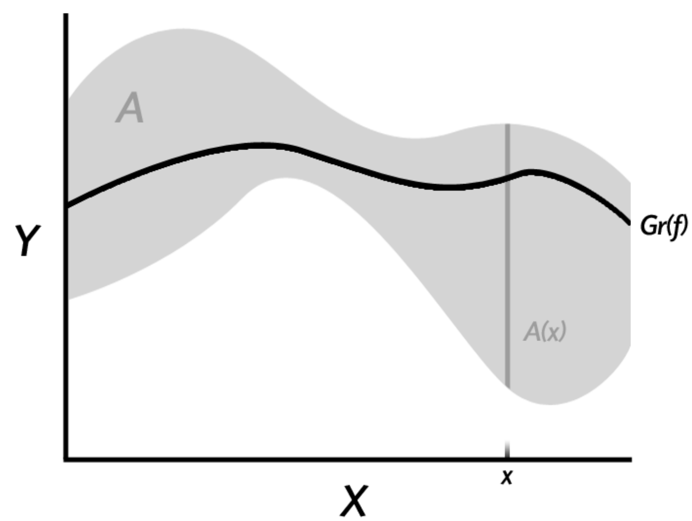

Figure 1: A choice set and a selection function.

tice, the existence of an arbitrary selection alone is often insufficient. In particular, the Axiom of Choice is agnostic to regularity in the parametrization of A: perhaps there is some relationship between $A(x)$ for different x that we want the selection to respect. Thus, it is typically preferable to have a selection satisfying some regularity property as a function, such as smoothness, continuity, or measurability (or to have a uniformization satisfying some regularity property as a set; this touches upon the issue of connecting the regularity of a function to the regularity of its graph, which we will return to later).

The theory of selection can then be seen as providing a “finer” or “more effective” version of the Axiom of Choice, where the central question addressed is: given a desirable property for f, what conditions on A guarantee the existence of a selection having that property? Of interest to us is the theory of measurable selection, which studies this question where the desirable property for f is measurability when X and Y are measurable spaces. Problems fitting into the selection framework arise in many ways; it will be useful to have a few simple examples in mind as we describe the surrounding machinery.

Example 1.1.1. One of the more basic selection problems is that of finding a left inverse to a surjective function $f : X \twoheadrightarrow Y$. This amounts to finding $g : Y \to X$ that is a selection from the graph $\mathrm{Gr}(f)$, viewed “sideways” as a subset of $Y \times X$. Regularity properties on f as a function translate to regularity properties on $\mathrm{Gr}(f)$, e.g. when X and Y are topological spaces, continuity of f implies that $\mathrm{Gr}(f)$ is closed (we will later discuss versions of this fact assuming measurability conditions instead of continuity for f). Thus, showing that regularity of $\mathrm{Gr}(f)$ implies the existence of a regular selection would give us a corresponding left inverse existence theorem (the inverse regularity is rarely as strong as we might wish; the most basic example of such a limitation is that continuous functions need not have continuous left inverses).

Example 1.1.2. In game theory, a player’s strategy can be represented as a selection from a collection of sets of available actions, each in response to some signal such as a history of previous turns. In this case, the external signal is coded in X, and the possible actions in Y. A strategy is a choice of action in response to all possible ensembles of external information, and thus a selection on some subset of $X \times Y$ (we constrain this subset somehow, often by some optimality condition that the strategy must attain—for instance, there may be a utility function on $X \times Y$ that we want to maximize, perhaps within some margin of error, as in the setting of the exact selection theorem in [4], which we will treat later).

1.2 The Metamathematical Simplification

Proving the existence of a measurable or otherwise regular selection function frequently involves fairly elaborate arguments that take extensive advantage of the particular structure of the sets involved. For example, we will look later at constructing, given a function $f : A \to \mathbb{R}$ on the choice set A, measurable selections of A whose points come close to minimizing or maximizing f on each $A(x)$ (the exact selection theorem mentioned above). It seems *a priori* that constructing such a selection would require working extensively with the metric properties of the $A(x)$ and the function

f, and this is indeed the approach typically taken in practice. And, measurable selection theorems often arise even more heavily contextualized by the application they are required for, in which case it is common for their proofs to use technical results and methods from the field of application to establish, say, the measurability of a selection function by establishing the measurability of certain inverse image sets, which may themselves have some independent domain-specific significance.

This sort of approach is effective but somewhat unsatisfying, for two primary reasons: firstly, modern descriptive set theory is typically conducted in the setting of *Polish spaces*, separable complete metric spaces whose metric has been "forgotten", the general accumulated philosophy of the field dictating that it is the topology of a Polish space that we should be working with as much as possible, not any particular concrete metrics that realize that topology. And, perhaps more importantly, results in modern set theory suggest that Lebesgue measurability (and the closely-related notion of universal or absolute measurability which we will work with later) of constructively defined sets encountered "in the wild" is, with the aid of enough extra foundational assumptions, something of a non-issue. The title of a well-known paper of Shelah and Woodin expresses this sentiment concisely: *Large cardinals imply that every reasonably definable set of reals is Lebesgue measurable* [21].

What this body of results means in practical terms is that when one encounters a set that is constructively defined (i.e. without the Axiom of Choice but rather say with just first-order quantifications), then if one is willing to assume the existence of sufficiently large cardinals, there is no need to do anything more to show that this set is measurable (in the Lebesgue or universal sense)—the very fact that the set is constructively defined ensures that measurability already.

Our goal here is to demonstrate that the same sort of simplification is possible for claims about the existence of measurable selections: we will see that the work of Shelah and Woodin along with some other metamathematical tools essentially imply that, assuming the existence of certain large cardinals, every "reasonably definable" planar subset of a Polish space has a Lebesgue measurable (or universally measurable) selection. We will illustrate by way of example how many results on measurable selections (or slightly weakened versions of these results) can be proven more easily and in a more uniform and systematic fashion by making these large cardinal assumptions and invoking some additional powerful results from set theory. We will also see that in practice, a relatively modest large cardinal assumption suffices for most of these arguments, namely the existence of a measurable cardinal.

1.3 Aside: On Interfacing Analysis and Descriptive Set Theory

There is a difficulty one inevitably encounters in trying to provide a lucid exposition of our subject matter, which is that the two kinds of objects whose relationship we are interested in—measurable spaces and measurable functions on them in analysis on the one hand, and large cardinals in set theory on the other—have a great deal of common theoretical ancestry, but are very far apart indeed on the map of modern mathematics as the average practitioner would draw it. The standard deftly-executed exposition of measure theory, for instance, gets by with a brief mention of the Axiom of Choice; from the perspective of introductory set theory, the Borel sets can be made to seem like their relevance consists purely in demarcating a natural complexity bound on subsets of infinite binary strings. That upon abandoning the Axiom of Choice one can find models of ZF in which every subset of $\mathbb{R}$ is Lebesgue measurable is a standalone folkloric trivium as far as practical analysis is concerned; in the set-theoretic context, "continuum" is a cardinal number devoid of geometric significance, and Lebesgue measurability in descriptive set theory is effectively little more than a regularity benchmark for subsets of $\mathbb{R}$ (which is why it is readily interchangeable in most set-theoretic results concerning it with many other properties, such as the perfect set property, the Baire property, universal measurability, and so forth).

This divide is unfortunate for us, because we are interested in precisely those theoretical domains where higher set theory and analysis can be observed causally jostling each other. Such cases are in fact plentiful, but the bearing of descriptive set theory on a given analytic problem is often quite opaque at first. Responsible for this are mostly linguistic and traditional differences: logicians prefer to talk about selections in terms of uniformizations (i.e. about the complexity of a selection function's graph as a set instead of the selection function as a mapping) while applications in analysis almost always framed in terms of selection functions, the machinery of descriptive set theory in practice analyzes the complexity of a definition in terms of the quantifiers involved (and economical use of quantifiers in writing a definition can be of the essence in obtaining effective complexity bounds) while in analysis quantifiers are little more than shorthand, etc. Some of our time will thus be spent stolidly translating between these two languages. To some extent though, that each problem we consider can be reduced to a rote calculation of complexities is the very quality of these problems that is on display here: we wish to demonstrate that a large class of analytic selection problems can basically be trivialized (or at least reduced to a mechanical and prescribed verification process) by a set-theoretic assumption.

Out project could thus reasonably be viewed as belonging to an undercurrent in modern mathematical logic towards finding "natural" or "concrete" statements in ordinary mathematics (i.e. statements having nothing to do with set theory or Gödel numbers of propositions or anything of that sort) whose truth still depends on assumptions independent of ZFC, such as large cardinals. (This mission statement is perhaps championed most vocally by Harvey Friedman, see e.g. [9] and [10].) Those statements, however, most often belong to combinatorics or number theory. Part of our motivation here is to contribute (minorly) to this program through problems with an analytic rather than a discrete flavor. Our contribution, it should also be noted, is in a slightly different spirit—instead of finding statements whose truth is contingent on independent assumptions, we exhibit statements whose ease of proof has this same contingency (this is, of course, still firmly in the vein of demonstrating that large cardinal axioms have concrete and practical ramifications).

2 Background

2.1 Complexity Hierarchies in Polish Spaces

As mentioned above, we pursue the question of measurable selection in the modern context of Polish spaces. The Polish spaces that will occur most often (and to which measurable selection theorems, particularly older results, will frequently be explicitly restricted, rather than generalizing to arbitrary Polish spaces) are $\mathbb{R}$ with its usual topology, ω with the discrete topology, ω^ω (the Baire space) with the product topology arising from the discrete topology on ω, and 2^ω (the Cantor space) with the product topology arising from the discrete topology on $2 = \{0, 1\}$. Another important example that will be useful occasionally is the space $C(X)$ of continuous functions $f : X \to Y$ where X is a compact Polish space and Y is an arbitrary Polish space, endowed with the supremum metric, $d_{C(X)}(f, g) = \sup_{x \in X} d_Y(f(x), g(x))$ (it is an easy exercise to verify that this is indeed again a Polish space). A final basic key fact is that countable products of Polish spaces are also Polish; hence, in particular, countable products of various collections of the above spaces are also Polish.

We will make use of the notion of pointclasses in Polish spaces as defined in [19], and we will adopt the following additional notations:

- For Γ a pointclass and X a Polish space, $\Gamma(X)$ denotes the subsets of X belonging to Γ.
- $\neg\Gamma$ denotes the pointclass of sets whose complements belong to Γ.
- Unions, intersections, and subset relations on pointclasses are defined in the obvious way.

The pointclasses we will be primarily concerned with are the Borel pointclasses Σ^0_α, Π^0_α, and Δ^0_α, and the projective pointclasses Σ^1_α, Π^1_α, and Δ^1_α where α ranges over countable ordinals (the definitions of these standard notations can be found in [19]). We also write $\mathcal{B}$ for the pointclass of all Borel sets (by a classical result equal to Δ^1_1). The notation of quantification over Polish spaces for sets and pointclasses will also be useful.

Definition 2.1.1. For X, Y Polish spaces and $A \subseteq X \times Y$, we write

$$\exists^Y A = \{x \in X : (\exists y \in Y)((x, y) \in A)\} = \pi_X(A),$$
$$\forall^Y A = \{x \in X : (\forall y \in Y)((x, y) \in A)\} = (\exists^Y A^c)^c.$$

For a pointclass Γ, we write

$$\begin{aligned}\exists^Y\Gamma &= \{\exists^Y A : A \in \Gamma(X \times Y) \text{ for a Polish space } X\} \\ &= \{\pi_X(A) : X \text{ a Polish space}, A \in \Gamma(X \times Y)\}, \\ \forall^Y\Gamma &= \{\forall^Y A : A \in \Gamma(X \times Y) \text{ for a Polish space } X\} \\ &= \neg(\exists^Y \neg\Gamma).\end{aligned}$$

Quantification over ω is closely related to the structure of the Borel hierarchy, and quantification over ω^ω is closely related to the structure of the projective hierarchy, as illustrated by the following structural results.

Proposition 2.1.2. For all $0 < \alpha < \omega_1$:

- Σ^0_α is closed under $\exists^\omega$ and Π^0_α is closed under $\forall^\omega$.
- $\exists^\omega\Pi^0_\alpha = \Sigma^0_{\alpha+1}$ and $\forall^\omega\Sigma^0_\alpha = \Pi^0_{\alpha+1}$.

Proposition 2.1.3. For all $0 \le \alpha < \omega_1$:

- Σ^1_α is closed under $\exists^{\omega^\omega}$ and Π^1_α is closed under $\forall^{\omega^\omega}$.
- $\Sigma^1_{\alpha+1} = \exists^{\omega^\omega}\Pi^1_\alpha$ and $\Pi^1_{\alpha+1} = \neg\Sigma^1_{\alpha+1} = \forall^{\omega^\omega}\Sigma^1_\alpha$.

These facts are useful for determining the descriptive complexity of subsets of Polish spaces in practice, where the defining property of a subset can typically be written in terms of such quantifiers over other Polish spaces, which allows the Borel or projective complexity of the subset to be read directly from its definition. An exposition of this perspective can be found in the first section of [19]. We will liberally blur the distinction between pointclasses and the definability properties that they represent, using terminology such as "a Borel condition" or "a Π^1_1 condition" to indicate the complexity of the set of points that satisfy the condition in question.

2.2 Heuristic Digression: Complexity vs. Regularity

Our methods largely rely on the possibility of exploiting the relationship between two types of "simplicity" properties or conditions for subsets of Polish spaces, and it will be useful to establish this distinction in heuristic terms before proceeding. Let X be a Polish space.

We call *complexity* conditions on a subset $A \subseteq X$ those conditions that bound where A is located in one of the above hierarchies (Borel or projective), which in practice corresponds to the quantifier complexity of a proposition ϕ whose truth determines whether a point belongs to A or not. A salient feature of complexity conditions is that they are unchanged by any reasonable additional metamathematical assumptions: the location of A in the boldface hierarchies depends exclusively on its definition and the formal structure of the proposition ϕ.

We call *regularity* conditions on $A \subseteq X$ those conditions such as Lebesgue measurability, universal measurability (defined in the following section), the Baire property, the perfect set property, and so forth, which make claims of the form "A contains a nice set" or "A is almost a nice set" for some interpretations of "almost" and "nice". What is relevant for our purposes about these conditions is that they are very much predicated on metamathematical assumptions: enough large cardinals or enough projective determinacy will make any projective A have all of the above properties, and the full Axiom of Determinacy will make any A at all (not necessarily projective) have all of the above properties, while the Axiom of Choice implies that there do exist subsets A of Polish spaces not having these properties (by constructions such as that of the Vitali sets).

Regular sets are not restricted in complexity—sets with any of the above regularity properties can be found at any level of the projective hierarchy. On the other hand, irregular sets *are* restricted in complexity, and the severity of the restriction depends on what axioms we use beyond ZFC: we will always have some result of the form that there are no irregular sets in Γ^1_α for all $\alpha < \alpha_0$ (for Γ either Σ or Π), and the bound α_0 will grow proportionally to the strength of the large cardinal assumption that we make. Hence, our approach, in broad strokes, is to take measurable selection theorems that impose complexity constraints on their measurable selections (i.e. complexity constraints on $f^{-1}(B)$ where f is the selection and B is Borel), to weaken these to instead make reference to regularity conditions, and then to ensure the satisfaction of these regularity conditions with a large cardinal assumption.

2.3 Notions of Measurability

As the preceding discussion indicates, we will deal with two fundamentally different types of measurability: measurability with respect to some projective pointclass Γ, and measurability with respect to some class of regular sets such as the Lebesgue measurable or universally measurable sets. We introduce these ideas in this section.

2.3.1 The Projective σ-Algebras

Measurability with respect to projective complexity classes requires us to consider the σ-algebra generated by a projective pointclass Γ, which we denote by the standard $\sigma(\Gamma)$. Note that measurability (Lebesgue or universal) for Γ implies the same measurability for $\sigma(\Gamma)$, so such measurability results will extend to the corresponding σ-algebras immediately.

We briefly establish some facts about the σ-algebras generated by the projective pointclasses: recall that Σ^1_α and Π^1_α are both closed under countable union and intersection, thus all that prevents them from already themselves being σ-algebras is closure under complementation. Since these two pointclasses consist of one another's complements, we clearly have $\sigma(\Sigma^1_\alpha) = \sigma(\Pi^1_\alpha)$. Let us denote this σ-algebra by σ^1_α. The pointclass Δ^1_α, on the other hand, is closed under countable unions and intersections as well as complements, so it already forms a σ-algebra. Thus, the projective σ-algebras consist of the σ^1_α and the Δ^1_α.

Frequently, the more general measurable selection theorems will be formulated over the domain of a general measurable space, i.e. a set Ω equipped with an arbitrary σ-algebra. We cannot handle this level of generality in the context of Polish spaces, so we will define here a suitable set of σ-algebras with which to replace such general conditions (in practice, it is rare that an application of one of these theorems is not over a Polish space).

Definition 2.3.1. A *Borel space* is the measurable space arising from a set Ω together with a σ-algebra $\mathcal{A}$ that is the collection of Borel sets generated by some topology on Ω (in the same way that we "forget" the metric of a Polish space, here we "forget" the topology of Ω and only remember its Borel sets). A *standard Borel space* or *standard $\mathcal{B}$ space* is a Borel space associated to some Polish space.

The above two definitions are standard in the theory of Polish spaces and Borel combinatorics, but we will find it useful to extend them to the larger σ-algebras defined above on Polish spaces: we let a *standard σ^1_α space* be any measurable space arising from a Polish space X equipped with the σ-algebra of $\sigma^1_\alpha(X)$.

When faced with a measurable selection theorem that makes reference to arbitrary measurable spaces, we will replace these with standard measurable spaces of some sort, usually $\mathcal{B}$ or σ^1_1 on some Polish space, to produce a version of the theorem in a Polish space setting (when this does not trivialize the theorem in question). We also mention in passing an important classical theorem that classifies the standard Borel spaces as measurable spaces by their cardinality.

Theorem 2.3.2 (Kuratowski). Any two standard Borel spaces of the same cardinality are Borel isomorphic. In particular, any standard Borel space is Borel isomorphic to either a discrete finite space, ω, or ω^ω (which is Borel isomorphic to $\mathbb{R}$).

2.3.2 The Lebesgue and Universal σ-Algebras

We will use the Lebesgue measure on any uncountable Polish space defined, say, using a fixed Borel isomorphism with $\mathbb{R}$ as per Kuratowski's theorem, and denote the collection of Lebesgue measurable sets in an uncountable Polish space X by $\mathcal{L}(X)$ (we will basically treat $\mathcal{L}$ as a pointclass akin to the projective pointclasses for the sake of notational consistency). We also often invoke a stronger notion of measurability for subsets of Polish spaces.

Definition 2.3.3. For X a Polish space, $A \subseteq X$ is *universally measurable* (*absolutely measurable* in some literature) if it is measurable with respect to every complete Borel probability measure on X. We denote by $\mathcal{U}(X)$ the collection of universally measurable sets in X, and, as above, treat $\mathcal{U}$ roughly as a pointclass.

We will see that large cardinal assumptions that imply the Lebesgue measurability of some pointclass also usually imply its universal measurability, and thus our results will often involve universal measurability as it is a stronger property: it is an easy exercise to show $\mathcal{U} \subsetneq \mathcal{L}$. And, it is a classical result that σ_1^1 sets are universally measurable (the result usually phrased as analytic sets being universally measurable), so we have a strict hierarchy $\mathcal{B} \subsetneq \sigma_1^1 \subsetneq \mathcal{U} \subsetneq \mathcal{L}$.

When it is not present explicitly in a measurable selection theorem already and the theorem demands a projectively measurable (i.e. Δ_α^1-measurable or σ_α^1-measurable) selection, we will substitute for this the condition of universal measurability. We will then produce a selection that is measurable with respect to some higher projective σ-algebra (for example, if the original theorem requires a σ_1^1-measurable selection, the results we leverage may only supply a Δ_2^1-measurable selection), and use a large cardinal assumption to force such a selection to also be universally measurable.

Remark. Note that moving from, say, Borel measurability to universal measurability is not, generally speaking, a particularly serious loss in strength for all practical purposes. In particular, a universally measurable function ϕ is equal μ-almost everywhere to a Borel measurable function, for any complete Borel probability measure μ. To see this, let $\mathcal{B}(\mu)$ be the completion of the Borel σ-algebra with respect to such a μ, then any universally measurable set is in $\mathcal{B}(\mu)$, whereby ϕ is $\mathcal{B}(\mu)$-measurable. But then, the definition of the completion $\mathcal{B}(\mu)$ implies that there is a map α_μ that is $\mathcal{B}$-measurable and equal to ϕ μ-almost everywhere. We will not bother with transforming our results into this form throughout, but this line of reasoning is a good intuitive indicator of the relative strength of universal measurability as a regularity condition for functions.

2.4 The Multifunction Viewpoint

Besides appealing to the visual intuition of drawing a curve through a planar set, measurable selection theory also often uses the multifunction formalism, which emphasizes that a selection makes a sequence of choices from sets.

Definition 2.4.1. A *multifunction* $F : X \rightrightarrows Y$ is a function with values in $\mathcal{P}(Y)$, the power set of Y (we take this notation to automatically designate that F is a multifunction).

Note that this is compatible with our notation of $A(x)$ for the section over x of some $A \subseteq X \times Y$; thus, a multifunction is to some extent identified with its graph. However, there are separate notions of measurability for multifunctions which do not always correspond directly with the same measurability for their graphs—we will deal with this obstacle below.

Given $F : X \rightrightarrows Y$, a selection for it is a function $f : X \to Y$ such that $f(x) \in F(x)$ for all $x \in X$. The graph $\mathrm{Gr}(F) \subseteq X \times Y$ of F is $\mathrm{Gr}(F) = \{(x, y) : y \in F(x)\}$.

Definition 2.4.2. We call $F : X \rightrightarrows Y$ *total* if $F(x) \neq \emptyset$ for all $x \in X$. For $F : X \rightrightarrows Y$ and a set $A \subseteq Y$, we define its *preimage*:

$$F^{-1}(A) = \{x \in X \mid F(x) \cap A \neq \emptyset\}.$$

For a σ-algebra $\mathcal{A}$ on X, we say that F is $\mathcal{A}$*-measurable* if $F^{-1}(U) \in \mathcal{A}$ for any open set $U \subseteq Y$, and *strongly* Γ*-measurable* if this holds for closed sets instead. A multifunction F is said to be Γ*-valued* for some collection of sets Γ if $F(x) \in \Gamma(Y)$ for each x (usually these are conditions like closed-valued or compact-valued). A brief disclaimer on notation is in order here: we adopt the definitions of Srivastava [23, Section 5.1] and others, but some sources [26] use *weakly* $\mathcal{A}$*-measurable* for our $\mathcal{A}$*-measurable*, and $\mathcal{A}$*-measurable* for our *strongly* $\mathcal{A}$*-measurable* (this is an earlier convention, after whose adoption many results indicated that weak measurability was actually the more fundamental notion, leading to the partial shift to our terminology).

It is easy to observe that when F is single-valued, then these notions both coincide with ordinary measurability. We note also that the conditions of being closed-valued or compact-valued, which often feature in multifunction selection theorems, do not force any particular complexity on the multifunction graph; for instance, any ordinary function is singleton-valued and hence closed-valued and compact-valued, but this does not impose any *a priori* restrictions on its graph. These are instead technical conditions that allow arguments such as those presented below, connecting a multifunction's measurability with the complexity of its graph. They may be thought of as imposing the condition that the multifunction be into some structured subset of $\mathcal{P}(Y)$, such as the Effros Borel structure on the closed subsets, or the (Polish) Vietoris topology on the compact subsets.

Proposition 2.4.3. Let X, Y be Polish spaces, and $F : X \rightrightarrows Y$ closed-valued. If F is Γ-measurable for a pointclass Γ that contains all of the closed sets and forms a σ-algebra, then $\mathrm{Gr}(F) \in \Gamma$.

Proof. Consider the condition $(x, y) \notin \mathrm{Gr}(F)$, which is equivalent to $y \notin F(x)$. Since $F(x)$ is closed, this means y is not a limit point of $F(x)$ and so if $\{U_n\}_{n\in\omega}$ is a countable basis for Y then for some n, we have $y \in U_n$ and $U_n \cap F(x) = \emptyset$. This is equivalently $x \notin F^{-1}(U_n)$, or $x \in F^{-1}(U_n)^c$. Thus, $(x, y) \notin \mathrm{Gr}(F)$ if and only if for some $n \in \omega$, $x \in F^{-1}(U_n)^c$ and $y \in U_n$. So, we find

$$\mathrm{Gr}(F)^c = \bigcup_{n\in\omega} F^{-1}(U_n)^c \times U_n,$$

or

$$\mathrm{Gr}(F) = \bigcap_{n\in\omega} (F^{-1}(U_n)^c \times U_n)^c = \bigcap_{n\in\omega} (F^{-1}(U_n) \times Y \cup X \times U_n^c).$$

Since F is Γ-measurable, $F^{-1}(U_n) \in \Gamma$, and so $F^{-1}(U_n) \times Y \in \Gamma$, and since Γ contains the closed sets, $X \times U_n^c \in \Gamma$ as well. Thus, every term of the intersection is in Γ, so $\mathrm{Gr}(F)$ is a countable intersection of sets in Γ, and thus is itself in Γ. □

In particular, this holds for any σ^1_α and any Δ^1_α. We can also move between these assumptions in the reverse direction.

Proposition 2.4.4. If Γ is a pointclass that contains the open sets and is closed under projection, and $F : X \rightrightarrows Y$ is a multifunction with $\mathrm{Gr}(F) \in \Gamma$, then F is $\sigma(\Gamma)$-measurable.

Proof. We have $F^{-1}(U) = \{x : F(x) \cap U \neq \emptyset\} = \pi_X(\mathrm{Gr}(F) \cap X \times U)$, which is in Γ if Γ is closed under projections. □

This in particular holds for any Σ^1_α with $\alpha \geq 1$. Note that since any ordinary single-valued function is a closed-valued multifunction, we automatically obtain from these results the convenient corresponding corollaries for single-valued functions as well, which will be useful later. We will soon examine some refinements of these results for membership of $\mathrm{Gr}(F)$ in the other projective pointclasses, as well.

3 Analytic Implications of Large Cardinals

Definition 3.0.5. The following table lays out the notation we will use for relevant set-theoretic hypotheses. The definitions and relevant background on any unfamiliar terms below, particularly those involving large cardinals, can be found in [12] or [13]. Let Γ and Γ_i below be collections of sets, and $n \leq \omega$.

Hypothesis	**Statement**
$\mathsf{Leb}(\Gamma)$	Any set in Γ is Lebesgue measurable.
$\mathsf{Univ}(\Gamma)$	Any set in Γ is universally measurable.
$\mathsf{Unif}(\Gamma)$	Any planar set in Γ has a selection whose graph is in Γ.
$\mathsf{MSel}_{\Gamma_1}(\Gamma_2)$	Any planar set in Γ_2 has a selection that is Γ_1-measurable.
$\mathsf{LebSel}(\Gamma)$	Any planar set in Γ has a Lebesgue measurable selection.
$\mathsf{UnivSel}(\Gamma)$	Any planar set in Γ has a universally measurable selection.
$\mathsf{Det}(\Gamma)$	The Gale-Stewart game on any set in Γ is determined.
PD	The Gale-Stewart Game on any projective set is determined.
MC	There is a measurable cardinal.
$\mathsf{Woo}(n)$	There are n Woodin cardinals with a measurable cardinal above all of them.

In this section, we present the existing results (and provide some interstitial propositions) linking the large cardinal hypotheses MC and $\mathsf{Woo}(n)$ to the measurable selection hypothesis LebSel, $\mathsf{UnivSel}$, and $\mathsf{MSel}_\bullet(\bullet)$. We will basically proceed backwards through the chain of relevant implications.

3.1 From Uniformization to Measurable Selection

The hypothesis $\mathsf{Unif}(\Gamma)$ above, uniformization of sets in Γ by functions with their graph in Γ, is used more commonly as an indicator of regularity for the pointclass Γ in set theory than hypotheses about measurable selections on Γ. However, these properties are closely related. Let $f : X \to Y$ be a map between Polish spaces, and let $\mathrm{Gr}(f) \subseteq X \times Y$ be its graph.

Proposition 3.1.1. If $\mathrm{Gr}(f) \in \Pi^1_n$, then f is Δ^1_{n+1}-measurable.

Proof. For a Borel set $B \subseteq Y$, we have $f^{-1}(B) = \pi_1((X \times B) \cap \mathrm{Gr}(f))$, where π_1 is projection onto the first coordinate. We have $X \times B$ is Borel, and $\mathrm{Gr}(f)$ is Π^1_n, so $f^{-1}(B)$ is a projection of a Π^1_n set and hence is Σ^1_{n+1}. But, the same is true for $f^{-1}(B^c) = f^{-1}(B)^c$, and so $f^{-1}(B) \in \Delta^1_{n+1}$. □

Proposition 3.1.2. If $\mathrm{Gr}(f) \in \Sigma^1_n$, then f is Δ^1_n-measurable.

Proof. We use the same argument as above, except taking advantage of the fact that Σ^1_n is closed under projection, thus $f^{-1}(B) \in \Sigma^1_n$ for any Borel $B \subseteq Y$. Again, the same holds for $f^{-1}(B)^c$, and so $f^{-1}(B) \in \Delta^1_n$. □

Corollary 3.1.3. The following implications hold.

- $\mathsf{Unif}(\Pi^1_n) \to \mathsf{MSel}_{\Delta^1_{n+1}}(\Pi^1_n)$.
- $\mathsf{Unif}(\Sigma^1_n) \to \mathsf{MSel}_{\Delta^1_n}(\Sigma^1_n)$.
- $\mathsf{Unif}(\Delta^1_n) \to \mathsf{MSel}_{\Delta^1_n}(\Delta^1_n)$.

Clearly we also have the two implications

- $\mathsf{MSel}_{\Gamma_1}(\Gamma_2) + \mathsf{Leb}(\Gamma_2) \to \mathsf{LebSel}(\Gamma_2)$,
- $\mathsf{MSel}_{\Gamma_1}(\Gamma_2) + \mathsf{Univ}(\Gamma_2) \to \mathsf{UnivSel}(\Gamma_2)$,

and in combination with the previous corollary, we obtain the following (we drop the results on Lebesgue measurability from this point onward because we will not need them later; from the foundational viewpoint universal measurability almost always comes along with Lebesgue measurability, so for the sake of economy we restrict ourselves to the stronger notion).

Corollary 3.1.4. The following implications hold.

- $\mathsf{Unif}(\Pi^1_n) + \mathsf{Univ}(\Delta^1_{n+1}) \to \mathsf{UnivSel}(\Pi^1_n)$.

- $\mathsf{Unif}(\Sigma^1_n) + \mathsf{Univ}(\Delta^1_n) \to \mathsf{UnivSel}(\Sigma^1_n)$.
- $\mathsf{Unif}(\Delta^1_n) + \mathsf{Univ}(\Delta^1_n) \to \mathsf{UnivSel}(\Delta^1_n)$.

3.2 Satisfying Uniformization

It remains to see how we can simplify the satisfaction of the hypotheses $\mathsf{Unif}(\Gamma)$ and $\mathsf{Univ}(\Gamma)$ for the projective pointclasses. First, we note that there are two classical theorems that give us a useful degree of uniformization in ZFC even without additional set-theoretic assumptions.

Theorem 3.2.1 (Novikoff-Kondô-Addison). $\mathsf{Unif}(\Pi^1_1)$ holds in ZFC.

Theorem 3.2.2 (Kondô). $\mathsf{Unif}(\Sigma^1_2)$ holds in ZFC.

We will refer liberally to both of these results as Kondô's theorem. Combining these statements with 3.1.4 gives us the following.

Corollary 3.2.3. $\mathsf{Univ}(\Delta^1_2) \to \mathsf{UnivSel}(\Sigma^1_2)$ (which subsumes $\mathsf{UnivSel}(\Pi^1_1)$).

This will largely suffice for our purposes, but it is worth mentioning the natural extension of Kondô's results by Moschovakis using the theory of scales, which pushes these results up the projective hierarchy assuming more projective determinacy.

Theorem 3.2.4 (Moschovakis 1971). $\mathsf{Det}(\Delta^1_{2n}) \to \mathsf{Unif}(\Pi^1_{2n+1})$ and $\mathsf{Unif}(\Sigma^1_{2n+2})$.

Note that at $n = 0$, we recover Kondô's theorem(s) from above (interpreting Δ^1_0 as the pointclass of clopen sets).

3.3 Satisfying Universal Measurability

Here we proceed by way of determinacy as an intermediate hypothesis between large cardinals and measurability of sets. We compose the following two well-known results.

Theorem 3.3.1 (Martin, Steel). For any $n \in \omega$, $\mathsf{Woo}(n) \to \mathsf{Det}(\Pi^1_{n+1})$.

Theorem 3.3.2 (Kechris, Martin). For any $n \in \omega$, $\mathsf{Det}(\Pi^1_n) \to \mathsf{Univ}(\Sigma^1_{n+1})$.

With this, we obtain the following.

Corollary 3.3.3. For any $n \in \omega$, we have that $\mathsf{Woo}(n) \to \mathsf{Univ}(\Sigma^1_{n+2})$, and thus also that $\mathsf{Woo}(n) \to \mathsf{Univ}(\sigma^1_{n+2})$.

In particular, at $n = 0$ we have $\mathsf{MC} \to \mathsf{Univ}(\sigma^1_2)$. Restricting this further, we obtain $\mathsf{MC} \to \mathsf{Univ}(\Delta^1_2)$, which in combination with Corollary 3.2.3 gives the following.

Corollary 3.3.4. $\mathsf{MC} \to \mathsf{UnivSel}(\Sigma^1_2)$.

This gives one method of extracting a measurable selection with the assistance of a measurable cardinal: we simply show that the set we are selecting from is Σ^1_2, which will ensure a universally measurable selection, the underlying situation being that we first produce a Σ^1_2-measurable uniformization, which is Δ^1_2-measurable as a function, and we then ensure that the Δ^1_2 sets are universally measurable.

3.4 Extending the Power of MC: A More Refined Argument

Note that in the above line of reasoning, we are wasting much of the power of the statement $\mathsf{MC} \to \mathsf{Univ}(\sigma^1_2)$ by only using that $\mathsf{Univ}(\Delta^1_2)$. Instead, it is sometimes advantageous to produce a uniformization partly by hand instead of just using the truth of $\mathsf{Unif}(\Gamma)$ for some Γ (we give an example below). Then, we can use the fact that MC implies that σ^1_2-measurable functions are universally measurable. One useful application of this line of reasoning is given below.

Example 3.4.1 (Piecewise Selections). Suppose X and Y are Polish spaces. Suppose also that we have $\{X_n\}_{n\in\omega}$ a partition of X into countably many disjoint subsets, with $X_n \in \sigma^1_2(X)$, and $A_n \subseteq X \times Y$ for $n \in \omega$, such that $X_n \subseteq \pi_X(A_n)$. The question we consider is whether we can produce a selection $f : X \to Y$ on which we impose the condition that $(x, f(x)) \in A_n$ whenever $x \in X_n$; that is, the set A_n represents some condition on f that it must satisfy over X_n.

Suppose that we can produce for each $n \in \omega$ a function $f_n : X \to Y$ that is a σ^1_2-measurable selection of A_n. Then, we can define $f : X \to Y$ piecewise by $f|_{X_n} = f_n$. For $R \subseteq Y$ an open set, we have $f^{-1}(R) = \bigcup_{n\in\omega} f_n^{-1}(R) \cap X_n$ and $f_n^{-1}(R)$ and X_n are both σ^1_2, so $f^{-1}(R) \in \sigma^1_2$, and thus f is σ^1_2-measurable. And, assuming a measurable cardinal, we have that f is universally measurable.

Various slightly stronger versions of this are possible. For instance, we can drop the requirement that the union of the X_n is all of X by just adding another subset $X_0 = (X_1 \cup X_2 \cup \dots)^c$ with $A_0 = X \times Y$ (i.e. on the area not covered by any X_n, we allow an arbitrary choice). We can also drop the requirement that the X_n are disjoint if there are only finitely many sets, by just enumerating all possible subcollections of the X_n that a given point could belong to (and intersecting the corresponding A_n).

3.5 Metamathematical Miscellanea

There are a number of metamathematical alternatives to assuming the existence of a measurable cardinal that can give us similar results; though we will work with a measurable cardinal in the following sections on applications of the preceding material, it is worth mentioning these options in case sharper assumptions are desirable.

3.5.1 Martin's Axiom and $\neg$CH

The first possible alternative is expressed in terms of the continuum hypothesis (CH) and Martin's axiom (MA; see [16] for the original description). We summarize in brief the story told in more detail by Martin and Solovay in [16]: it was discovered in short order after Cohen's proof of the independence of CH that $\neg$CH was a much weaker statement than CH, failing to decide many of the interesting statements that could be proved with CH. Martin's axiom MA was of interest in part because MA + $\neg$CH is in fact of roughly commensurable logical strength with CH (though the network of relationships here is rather elaborate; we also have for instance CH $\to$ MA, and MA itself implies many of the interesting consequences of CH).

As described in the original paper (and as is now conventional notation), there is a parametrization by cardinal numbers of Martin's axiom, giving axioms $\mathsf{MA}(k)$ for a cardinal k, where MA itself is the statement that if $k < 2^{\aleph_0}$, then $\mathsf{MA}(k)$. It is shown in [16] that $\mathsf{MA}(\aleph_1)$ (which of course follows from MA + $\neg$CH) implies $\mathsf{Leb}(\Sigma^1_2)$ (and it is easy to see from the argument given there that the same reasoning gives universal measurability as well), which gives us much the same power as a measurable cardinal (using Kondô's theorem for Σ^1_2 to produce the uniformization before applying this).

A variety of combinations of assumptions are also possible and give a corresponding variety of possibly-relevant results; perhaps the most enticing is, as shown in [16], that $\mathsf{MA}(\aleph_2) + \mathsf{MC} \to \mathsf{Leb}(\Sigma^1_3)$, and correspondingly also $\mathsf{Univ}(\Sigma^1_3)$. However, to take advantage of this directly to obtain, say, $\mathsf{UnivSel}(\Sigma^1_3)$ would require a Woodin cardinal so that we can uniformize Σ^1_3 sets in the first place, and we will see that $\mathsf{UnivSel}(\Sigma^1_2)$ is sufficient for all examples we consider, so we will not pursue this line of reasoning further here.

3.5.2 Shoenfield Absoluteness and Absolute Descriptive Complexity

Fenstad and Normann have also introduced notions of *provable* and *absolute* descriptive complexity, particularly looking at these versions of the Δ^1_2 pointclass [20].

Definition 3.5.1. A set $A \subseteq \omega^\omega$ is *provably* Δ^1_2 if there are Σ^1_2 and Π^1_2 propositions Ψ and Φ, respectively, with free variables x and y, and a parameter $y \in \omega^\omega$ such that:

- $x \in A \leftrightarrow \Psi(x, y) \leftrightarrow \Phi(x, y)$, and
- $\mathsf{ZF} \vdash \forall(x, y \in \omega^\omega)(\Psi(x, y) \leftrightarrow \Phi(x, y))$.

And, A is *absolutely* Δ^1_2 if we have such propositions Ψ, Φ again such that:

- $x \in A \leftrightarrow \Psi(x, y) \leftrightarrow \Phi(x, y)$, and
- for all standard models M of ZF such that $x, y \in M$, we have $\Phi(x, y) \leftrightarrow M \vDash \Phi(x, y)$ and $\Psi(x, y) \leftrightarrow M \vDash \Psi(x, y)$.

We note, as do the authors of [20], that it is immediate from Shoenfield's absoluteness theorem that if A is provably Δ^1_2 then it is also absolutely Δ^1_2. The main result proved on these sets is the following.

Theorem 3.5.2 (Fenstad, Normann). If for all $x \in \omega^\omega$ there is a countable standard model M of ZF such that $x \in M$, then every absolutely Δ^1_2 set is universally measurable.

Thus, so long as we can obtain selections ϕ such that preimages of open/Borel sets are not just Δ^1_2 but are provably so, then we can reproduce these results without any metamathematical assumptions, avoiding even the need to assume a measurable cardinal.

In particular, with reference to an "effective" proof of Kondô's theorem, with Π^1_1 (or Σ^1_2 in the other version) taken as arithmetical hierarchy complexities in ω^ω relative to a fixed element $a \in \omega^\omega$—historically, an effective proof was first provided by Addison, which proof is concisely presented in [13, Theorem 13.17]—we see as a consequence that any Δ^1_2-measurable selection produced by Kondô's theorem can be chosen to be provably Δ^1_2, and thus to be universally measurable by the above (alternatively, in all of the below proofs we of course also determine the complexity of the preimage sets for a selection function by hand, thus implicitly demonstrating that these sets are provably Δ^1_2 in each case individually also).

Perhaps of further independent interest from the same work may be their note that the method of the proof of the above theorem, in combination with the Shoenfield absoluteness theorem and the MC assumption, gives $\mathsf{Univ}(\Sigma^1_2)$ as well, according to an unpublished argument of Solovay's. This gives a much more direct derivation of $\mathsf{MC} \to \mathsf{Univ}(\Sigma^1_2)$ with significantly less high-powered set-theoretic machinery involved than the path that we took above (though our approach has the advantage of illustrating the extensions of these statements up the projective hierarchy under Woodin cardinal assumptions).

4 Applications

> *My own view is that the projective hierarchy is so far removed from normal mathematics that...this development does not come close to satisfactorily establishing the importance or relevance of large cardinals to mathematics. We want demonstrably necessary uses of large cardinals in concrete mathematical contexts—the more concrete the better.*
>
> — Harvey Friedman, *What You Cannot Prove*

In this section, we provide some illustrations of actual applications of measurable selection ideas in relatively "concrete" mathematics, mostly either mathematical economics or the related areas of optimization and control theory (in rather abstractlyconceived form). First though, we will examine how our results compare in strength to some of the classical results from the general theory of measurable selection.

4.1 General Selection Theorems

In this section we examine some of the historically original theorems of measurable selection theory, which typically attempted to produce selections of a certain measurability under a particular assumption of the complexity of the choice set, without assuming any further structure on the set (indeed note that Kondô's theorem is such a statement; we use it among our tools as it is weaker and chronologically prior by roughly a decade to the body of work under consideration here).

4.1.1 Selections of Borel and Analytic Sets

One of the initial motivating questions of measurable selection theory was: under what conditions does a planar Borel set have a Borel-measurable selection? The answer cannot be "always", as illustrated by a standard counterexample due to Blackwell [23, Example 5.1.7]. This obstacle can be circumvented in one of two directions: either one can find how much stronger the assumption on the Borel set must be in order to force it to have a Borel selection, or one can weaken the assumption on the Borel-measurable selection in the conclusion until the statement is true.

Only the second of these questions is within the reach of our methods: Borel sets are the smallest σ-algebra that we are able to distinguish in the context of the projective hierarchy, and we recall that our plan consisted in replacing stronger notions of measurability with universal measurability to render large cardinal assumptions useful. One of the well-known results following this route is the following theorem (attributed variously to Jankoff and von Neumann), which actually proves a more general statement about selecting from analytic sets.

Theorem 4.1.1 (Jankoff-von Neumann [25]). Let X, Y be Polish spaces and $F : X \rightrightarrows Y$ a total closed-valued multifunction with $\mathrm{Gr}(F) \in \Sigma^1_1(X \times Y)$. Then, F admits a σ^1_1-measurable selection.

If we weaken the conclusion of this theorem to universal measurability, then we immediately obtain a proof assuming a measurable cardinal from $\mathsf{MC} \to \mathsf{UnivSel}(\Sigma^1_2)$, since $\sigma^1_1 \subseteq \Sigma^1_2$, and indeed if we do this then we may also further weaken the assumption on the graph of F to $\mathrm{Gr}(F) \in \Sigma^1_2(X \times Y)$.

4.1.2 The Kuratowski-Ryll Nardzewski Theorem

It is worth mentioning one of the main broad generalizations of von Neumann's theorem, though we will only be able to prove a relatively narrow restriction of its claim by our means. We present a slight restriction of the theorem in its full generality;

our version is the one discussed in Wagner's survey [26]. A standard proof of this version can be found in [1].

Theorem 4.1.2 (Kuratowski-Ryll Nardzewski). If $(S, \mathcal{A})$ is a measurable space, Y a Polish space, and $F : (S, \mathcal{A}) \rightrightarrows Y$ a total, measurable multifunction with closed values, then F admits an $\mathcal{A}$-measurable selection.

To adapt this to proof by our means, we replace the domain with another Polish space X, and replace the conclusion with F having a universally measurable selection. Since we want to use $\mathsf{MC} \to \mathsf{UnivSel}(\Sigma^1_2)$, the largest projective σ-algebra on X that we can handle here is Δ^1_2. Then, we can prove the following variation of the theorem (this is something of a triviality relative to the power of the original theorem, but it is another useful basic illustration of our approach).

Theorem 4.1.3 (Δ^1_2 Kuratowski-Ryll Nardzewski). If X and Y are Polish spaces, and $F : X \rightrightarrows Y$ a total, Δ^1_2-measurable multifunction with closed values, then F admits a universally measurable selection.

Proof (with MC*).* We use our earlier result that a Δ^1_2-measurable multifunction has a Δ^1_2 graph, and thus in particular a Σ^1_2 graph, and invoking $\mathsf{MC} \to \mathsf{UnivSel}(\Sigma^1_2)$ then completes the proof. □

4.1.3 Countable Families of Selections

It is also common that a measurable selection theorem produce not just one measurable selection, but a countable sequence of measurable selections, which must cover (under various interpretations of "covering") the choice set. Fortunately, this setting is not so different from the case of a single selection, as our approach to the first result below will illustrate.

The simplest setting of the above sort of theorem is when the sections of the choice set are countable, in which case it is reasonable to expect that countably many selections can cover the entire choice set. The theorem below is the most basic result of this kind.

Theorem 4.1.4 (Lusin). Let X, Y be Polish spaces and $A \in \Sigma^1_1(X \times Y)$ such that $A(x)$ is countable for all $x \in X$. Then, there are countably many Borel selections $\phi_i : X \to Y$ of A for $i \in \omega$ such that $\bigcup \mathrm{Gr}(\phi_i) = A$, i.e. the graphs of the ϕ_i cover A.

One standard proof of this uses as a preliminary a reflection theorem due to Burgess, from which Lusin's result follows more or less directly [23, Theorem 5.10.3]. As before, we weaken the condition on the ϕ_i from Borel measurability to universal measurability before proceeding.

Proof (with MC*).* This proof will illustrate the general approach we will take to producing countable sequences of selections: instead of trying to produce the individual functions ϕ_i, we instead produce a map $\Phi : X \to Y^\omega$ (recall that Y^ω is still Polish), such that upon taking ϕ_i to be the ith coordinate of Φ, we obtain the desired sequence of ϕ_i. Note that composing a universally measurable function with the projection onto the ith coordinate gives a universally measurable function, since the projections are continuous. Thus, for the ϕ_i to be universally measurable, it suffices for Φ to be universally measurable.

The condition that the graphs of the ϕ_i cover A is equivalently:

$$\forall (x \in X) \forall (y \in Y) \Big((x, y) \in A \to \exists (i \in \omega)(\phi_i(x) = y) \Big),$$

or

$$\forall (x \in X) \forall (y \in Y) \exists (i \in \omega) \Big((x, y) \notin A \text{ or } \pi_i(\Phi(x)) = y \Big),$$

hence Φ must be a selection from the set $P \subseteq X \times Y^\omega$ of (x, z) that satisfy

$$\forall (y \in Y) \exists (i \in \omega) \Big((x, y) \notin A \text{ or } \pi_i(z) = y \Big).$$

Since A is Σ^1_1 and π_i is continuous for each i, the inner condition is Π^1_1, and hence the set P is Π^1_1 (and trivially has non-empty sections; over any x, just take z to enumerate $A(x)$). Hence, P has a universally measurable selection. □

A variant of this statement due to Burgess (presented in its source for the purposes of resolving a problem in mathematical economics) can be found in [5], where the large cardinal simplification is also mentioned briefly. We elaborate the argument in our notation here. The section of the argument that we simplify concerns the following result, Corollary 13 in [5].

Theorem 4.1.5 (Burgess). Let X, Y be Polish spaces, and $A \in \Sigma^1_1(X \times Y)$. Then:

- The set $C = \{x \in X : |A(x)| \leq \omega\}$, the set of x with countable sections, is Π^1_1.
- There is a sequence $\{\phi_i\}_{i \in \omega}$ of countably many σ^1_1-measurable selections $\phi_i : X \to Y$ such that for all $x \in C$, $A(x) \subseteq \{\phi_i(x) : i \in \omega\}$.

The measurable selection is the second part, and we will show how to prove the second part assuming the first. We reduce the σ^1_1-measurability condition to universal measurability, and proceed as before.

Proof (with MC*).* The proof follows the one for Lusin's theorem exactly, except that the condition

$$(x, y) \in A \to \exists (i \in \omega)(\pi_i(\Phi(x)) = y)$$

is only required when $x \in C$, which we can write as an extra condition

$$((x, y) \in A \text{ and } x \in C) \to \exists (i \in \omega)(\pi_i(\Phi(x)) = y),$$

and since $A \in \Sigma^1_1$ and $C \in \Pi^1_1$, their intersection is only guaranteed to be contained in Δ^1_2, thus the choice set P can only be assumed to be Δ^1_2. But, since our universal selection results under a measurable cardinal extend to Σ^1_2, this still has a universally measurable selection, and so the rest of the proof goes through as for Lusin's theorem above. □

Note that if desired, an intermediate result of this argument yields a family of Δ^1_2-measurable selections satisfying the necessary condition, a different weakening of the σ^1_1-measurability in the original result.

4.1.4 Castaing Representations

Now, we explore the selection theorems resulting from lifting the condition that the sections of the choice set we are concerned with be countable. The easiest way to adjust the types of statements examined above into this context is to require that the section above some x be covered not by just the union of the selection values $\{\phi_i(x) \mid i \in \omega\}$, but by the closure of this set.

Definition 4.1.6. For X, Y Polish spaces and a closed-valued multifunction $F : X \rightrightarrows Y$, we call a family of selections $\{\phi_i\}$ a *Castaing representation* of F if for each $x \in X$, $F(x) = \overline{\{\phi_i(x) \mid i \in \omega\}}$. A property demanded of the representation is demanded of each function, for instance a measurable Castaing representation has each ϕ_i measurable.

Theorem 4.1.7 (Castaing). Let $(X, \mathcal{A})$ be a measurable space, Y a Polish space, and $F : X \rightrightarrows Y$ a total, closed-valued multifunction. If F is $\mathcal{A}$-measurable, then F admits an $\mathcal{A}$-measurable Castaing representation.

This theorem and its standard proof can be found in [6, Theorem III.8.1] (indeed, a brief argument given in their Theorem III.9 shows that F being $\mathcal{A}$-measurable is equivalent to the existence of an $\mathcal{A}$-measurable Castaing representation, so Castaing representations provide an alternate characterization of multifunction measurability). We will prove this in lesser generality, where $(X, \mathcal{A})$ is a standard Borel space, on the Polish space X, and where the functions of the Castaing representation are universally measurable.

Proof (with MC*).* As before, reinterpret the required universally measurable $\phi_i : X \to Y$ to a single universally measurable $\Phi : X \to Y^\omega$. Let $\{U_n\}$ be a countable basis of Y. The condition that $\{\phi_i(x) \mid i \in \omega\}$ is dense in $F(x)$ can then be written as

$$\forall (r \in F(x)) \forall (n \in \omega) \Big(r \in U_n \to \exists (i \in \omega) (\pi_i(\Phi(x)) \in U_n) \Big),$$

and so Φ is a selection from the set P consisting of $(x, z) \in X \times Y^\omega$ for which

$$\forall (r \in F(x)) \forall (n \in \omega) \Big(r \in U_n \to \exists (i \in \omega)(\pi_i(z) \in U_n) \Big),$$

which after expanding the implication and extracting the quantifier over ω is

$$\forall (r \in Y) \forall (n \in \omega) \exists (i \in \omega) \Big(\neg((x, y) \in \mathrm{Gr}(F) \text{ and } r \in U_n) \text{ or } \pi_i(z) \in U_n \Big),$$

or

$$\forall (r \in Y) \forall (n \in \omega) \exists (i \in \omega) \Big((x, y) \notin \mathrm{Gr}(F) \text{ or } r \notin U_n \text{ or } \pi_i(z) \in U_n \Big).$$

Since F is Borel-measurable, $\mathrm{Gr}(F)$ is a Borel set, so the entire inner condition is Borel, as is the condition including both quantifications over ω. So, we have $P \in \Pi^1_1(X \times Y^\omega)$, and thus P has a universally measurable selection. □

4.2 Measurably Parametrized Theorems

Though it is not the case across the board, it is very typical for measurable selection theorems in more concrete fields to come in a particular flavor, namely that of a "randomization" or "parametrization" of an existing result.

Very broadly speaking, this consists of the following transformation of a result: suppose that given Polish X and Y, for any $x \in X$ we can construct $y \in Y$ to satisfy some proposition $P(x, y)$. Then, given a measurable space Ω, we ask if, given a measurable $x : \Omega \to X$, we can construct a measurable $y : \Omega \to Y$ such that for all $\omega \in \Omega$, $P(x(\omega), y(\omega))$ holds. In other words, we ask if the construction can be carried out in a measurable way with respect to a new parameter ω, assuming that the "problem statement" x behaves measurably with respect to this parameter. If Ω is a probability space, then x and y are random variables, in which case another interpretation is that there is uncertainty in the problem statement x, and we want a probabilistic construction y such that $P(x, y)$ still always holds.

4.2.1 A First Parametrization: The Bolzano-Weierstrass Theorem

We begin with a simple example to illustrate the general course of action, parametrizing a familiar and basic theorem from analysis. It is worth nothing that this is not just a novelty; the statement we work with here (though easy to prove directly) is used e.g. in [8] in a proof of the "fundamental theorem of asset pricing", a characterization of arbitrage-free stochastic financial market models.

The Bolzano-Weierstrass theorem states that any bounded sequence $\{x_n\}_{n \in \omega}$ in $\mathbb{R}^d$ has a convergent subsequence, i.e. a strictly increasing sequence $\{\sigma_m\}_{m \in \omega}$ in ω, which can also be interpreted as a single element $\sigma \in \omega^\omega$, and an element $y \in \mathbb{R}^d$ such that $x_{\sigma_m} \to y$ (we write the convergent-subsequence condition in such expanded form to make the parallel with the following version very transparent).

Suppose we now have a measurable space $(\Omega, \mathcal{F})$ against which we would like to parametrize this. We take $x_n : \Omega \to \mathbb{R}^d$ (all functions we mention from Ω will be measurable) such that $\{x_n(\omega)\}$ for any given ω is a bounded set. Then, we want to find $\sigma : \Omega \to \omega^\omega$ (whose values are restricted to the subset of strictly increasing sequences in ω^ω) and $y : \Omega \to \mathbb{R}^d$ such that $x_{\pi(\sigma(\omega),m)}(\omega) \to y(\omega)$ for all $\omega \in \Omega$, where $\pi(a,k)$ is the function projects a from whatever product space a belongs to, onto its kth coordinate. Write $X = (\mathbb{R}^d)^\omega$ (a Polish space), then we can view all the x_n together as a single map $x : \Omega \to X$, so that the convergent-subsequence condition becomes $\pi(x(\omega), \pi(\sigma(\omega), m)) \to y(\omega)$.

We then view the pair of desired functions (y, σ) as a single map $\phi : \Omega \to \mathbb{R}^d \times \omega^\omega$, and construct a choice set $C \subseteq \Omega \times \mathbb{R}^d \times \omega^\omega$ from which we will select ϕ. This choice set contains a point (ω, y, σ) if and only if σ is a strictly increasing sequence (which is a closed condition, as any limit point of a set of strictly increasing sequences must clearly itself be strictly increasing), and $\pi(x(\omega), \pi(\sigma, m)) \to y$, which latter condition can be rewritten as:

$$\forall(n \in \omega)\exists(N \in \omega)\forall(m \in \omega)\left(m \geq N \to |\pi(x(\omega), \pi(\sigma, m)) - y| < \frac{1}{n}\right).$$

This is a Borel condition, and hence C is a Borel set. Also, we have that C is total as a multifunction by the ordinary unparametrized Bolzano-Weierstrass theorem.

Now, suppose that Ω is a standard Borel space. Then, the inner condition is Borel, and thus C is Borel, and has a universally measurable selection, which gives a universally measurable ϕ as above (we could just as well assume that Ω is a standard σ_1^1 space, or in fact any sub-σ-algebra thereof; note though that since we are producing a universally measurable selection against the Polish space structure of Ω, not an $\mathcal{F}$-measurable selection, we may as well formulate the result in terms of the finest possible σ-algebra). This gives the desired parametrization of the Bolzano-Weierstrass theorem; we have proved the following (unpacked back into more familiar form).

Theorem 4.2.1. If Ω is a Polish space and $x_n : \Omega \to \mathbb{R}^d$ are σ_1^1-measurable, with each $x(\omega)$ as a sequence in $\mathbb{R}^d$ having a convergent subsequence, then there are universally measurable $\sigma_m : \Omega \to \omega$ and $y : \Omega \to \mathbb{R}^d$ such that for each $\omega \in \Omega$, $\sigma_m(\omega)$ is a strictly increasing sequence in m, and $x_{\sigma_m(\omega)}(\omega) \to y(\omega)$.

4.2.2 Parametrized Continuous Preference Orders

Of central interest in mathematical economics is the question of representing preference orders by utility functions. Here, we consider the question of parametrizing this problem.

Definition 4.2.2. For a topological space X, a *continuous preference order* on X is a binary relation $\succeq$ that is reflexive, transitive, and total, and so that for any $x \in X$, the preceding and succeeding segments, $\{x' \in X : x \succeq x'\}$ and $\{x' \in X : x' \succeq x\}$

respectively, are both closed sets in the topology of X. A continuous function $f : X \to \mathbb{R}$ *represents* $\succeq$ if we have $x \succeq x'$ if and only if $f(x) \geq f(x')$.

Of particular interest are continuous (as functions) representations of continuous preference orders. An older theorem of Debreu cited in [27] gives that a continuous preference order on a second-countable space X can be represented by a continuous function. The Measurable Utility Theorem proved by Wesley in [27] states a version of the following (which is adapted for our purposes).

Theorem 4.2.3 (Adapted Measurable Utility Theorem). Let T, X be Polish, let $\phi : T \rightrightarrows X$, and for each $t \in T$ let $\succeq_t$ be a continuous preference order on the set $\phi(t) \subseteq X$. Suppose that the set

$$S = \{(t, x, x') : x, x' \in \phi(t) \text{ and } x \succeq_t x'\} \subseteq T \times X \times X$$

is Borel, then there is a universally measurable function $U : T \times X \to \mathbb{R}$ so that for almost all $t \in T$, the function $U(\cdot, t)$ is continuous and represents $\succeq_t$.

In the original, we have $T = [0, 1]$, X a finite power of $\mathbb{R}$, the conclusion only holds for Lebesgue-almost all t, and the result a Borel measurable function instead of a universally measurable one (recall also from our earlier discussion that it is possible to parlay a universally-measurable function into Borel function that is μ-almost everywhere identical for any complete Borel measure μ, so we can easily recover something resembling the original result from our adapted version in this case).

Proof (with MC*).* Note that it is enough to prove the statement for X compact: any Polish space embeds as a G_δ subset of the Hilbert cube $[0, 1]^\omega$, which is compact, and clearly expanding the space X in the theorem does not change the statement (as ϕ is allowed to be an arbitrary multifunction).

Then, we instead will view U as a function $U : T \to C(X)$, which subsumes the requirement that $U(\cdot, t)$ be continuous, and where $C(X)$ is Polish since X is compact. Let $\Phi \subseteq T \times X$ be the graph of ϕ. Then, note that since $x \succeq_t x$ for each $x \in \phi(t)$ for any t, if we define the diagonal of S as $S_D = S \cap \{(t, x, x') : x = x'\}$ (which is again Borel), then Φ is the projection of S_D onto $T \times X$, whereby $\Phi \in \Sigma^1_1(T \times X)$.

Let $C \subseteq T \times C(X)$ be the choice set we will construct. We let $(t, f) \in C$ if and only if:

$$\forall (x, x' \in X)\Big((t, x) \in \Phi \text{ and } (t, x') \in \Phi \to \big(f(x) \geq f(x') \leftrightarrow (t, x, x') \in S\big)\Big).$$

Note that the inner condition is just that of f representing $\succeq_t$. Since S is Borel, Φ is Σ^1_1, and $f(x) \geq f(x')$ is a Borel condition in the variables involved, we have that the inner condition is at worst Δ^1_2, and thus the entire condition is at worst Σ^1_2, hence $C \in \Sigma^1_2(T \times C(X))$. And, C is total as a multifunction, since by Debreu's theorem mentioned above we have that each $\succeq_t$ is indeed represented by some continuous function on X. Thus, assuming a measurable cardinal, C has a measurable selection, which then satisfies the desired criteria. □

This is a particularly interesting application historically and metamathematically speaking, as the original result is proved in [27] using methods involving generic models and Cohen forcing, and at the time of writing the authors knew of no "elementary" method of proving the claim (i.e. one not making use of external logical tools), and as on the other hand the above method of proof of the adapted version does not at all suggest that this theorem should be significantly more difficult than many of the results that we considered earlier.

A somewhat more similar result to our adapted version is proved in [18] (the continuous representations of continuous preference orders are also called *Paretian utility functions*). In particular, this proves the result for all indices $t \in T$ rather than for Lebesgue-almost all of them. A further, more modern discussion of a similar problem can also be found in [11].

4.2.3 Parametrized Universalities of the Cantor Space

One of the earliest motivations for the development of descriptive set theory was Cantor's attempt to prove the truth of the Continuum Hypothesis (CH) by showing, for any $A \subseteq \mathbb{R}$, that either A is countable, or else A contains a homeomorphic image of the Cantor space 2^ω. In light of the independence of CH, this property is of course only true of particular classes of regular subsets of $\mathbb{R}$; in particular, a construction due to Cantor [19, Corollary 1A.3] gives that any non-empty perfect $P \subseteq \mathbb{R}$ contains inside it a homeomorph of 2^ω. For Cantor, the useful corollary of this was that any subset of $\mathbb{R}$ containing a non-empty perfect set then could not have a cardinality lower than that of the continuum (by the Cantor-Bendixson theorem, the closed sets of any Polish space have the so-called "perfect set property", a regularity property very commonly used in descriptive set theory of a subset either being countable or else containing a perfect set; another classical result gives the same for all Σ^1_1 sets; sufficiently large cardinals give the same for any projective set).

We examine here a parametrized version of the problem of finding a homeomorphism $2^\omega \hookrightarrow P$ for P a perfect subset of a Polish space (which by classical means usually entails directly parametrizing Cantor's construction mentioned above). We look at Pol's theorem [15] (this source also provides a generalization weakening σ^1_1-measurability to Borel measurability, which we do not consider separately because our methods require weakening both of these conditions to universal measurability anyway), which states the following.

Theorem 4.2.4. Let X be Polish, and $F : X \rightrightarrows \omega^\omega$ total with $\mathrm{Gr}(F) \in \Sigma^1_1$ such that $F(x)$ is a (non-empty) perfect subset of ω^ω for each $x \in X$. Then, there is a function $\phi : X \times 2^\omega \to \omega^\omega$ such that for each x, $\phi(x, \cdot)$ is a homeomorphism of 2^ω into $F(x)$, and for each $y \in 2^\omega$, $\phi(\cdot, y)$ is σ^1_1-measurable as a function $X \to \omega^\omega$.

We weaken the σ_1^1-measurability to universal measurability, and instead of having $\phi(\cdot, y)$ be measurable for each y, we view ϕ alternately as a map $X \to C(2^\omega)$, and force ϕ to be measurable as such a map (the map $C(2^\omega) \to \mathbb{R}$ given by evaluation at any $y \in 2^\omega$ is continuous, so if ϕ is universally measurable in this sense, then $\phi(\cdot, y)$ will also be universally measurable for each y; also, we use that ω^ω embeds into $\mathbb{R}$ as the set $\mathcal{I}$ of irrational numbers, so we can take an element $f \in C(2^\omega)$ as a continuous map $2^\omega \to \omega^\omega$ if $f(a) \in \mathcal{I}$ for all $a \in 2^\omega$).

Proof (with MC*).* Let us first establish a bound on the complexity of the subset $H \subseteq C(2^\omega)$ of continuous functions that are homeomorphisms onto their image (i.e. are one-to-one and have continuous inverses). The condition that $h \in C(2^\omega)$ be one-to-one is

$$\forall (x, y \in 2^\omega)\Big(x = y \text{ or } h(x) \neq h(y)\Big),$$

which is Π_1^1. The condition that there exist a continuous inverse j is

$$\exists (j \in C(2^\omega))\forall (x \in 2^\omega)\Big(j(f(x)) = x\Big),$$

which is Σ_2^1. Thus, $H \in \Sigma_2^1$.

We construct a choice set C for the function $\phi : X \to C(2^\omega)$. For each $x \in X$, we firstly require $\phi(x) \in H$ so that $\phi(x)$ is a homeomorphism. Let us fix a homeomorphism $\alpha : \mathcal{I} \to \omega^\omega$. Then, we require that $\operatorname{img}(\phi(x)) \subseteq \mathcal{I}$, and that $\alpha \circ \phi(x)$ is onto $F(x)$. In general, for $f \in C(2^\omega)$, the property of $\operatorname{img}(f) \subseteq \mathcal{I}$ is that $\forall (x \in 2^\omega)(f(x) \in \mathcal{I})$ which since $\mathcal{I}$ is Borel is a Π_1^1 condition.

And, the property that $\alpha \circ f$ be onto $F(x)$ is in the most immediate form

$$\forall (y \in \omega^\omega)\exists (z \in 2^\omega)\Big((x, y) \in F \to \alpha(f(z)) = y\Big),$$

but this would give us a complexity bound of Π_2^1 which is undesirable as a measurable cardinal does not in that case ensure a measurable selection. So, we instead fix a countable dense subset $\{z_i\}_{i\in\omega}$ of 2^ω, and write the equivalent condition

$$\forall (y \in \omega^\omega)\forall (n \in \omega)\exists (i \in \omega)\Big((x, y) \in F \to d(\alpha(f(z_i)), y) < \frac{1}{n}\Big),$$

where d is a fixed metric for Baire space. This is an equivalent condition because the image of a dense set under a continuous map (the $\alpha(f(z_i))$ in our case) is dense in the map's image. Now, we recall that $\operatorname{Gr}(F) \in \Sigma_1^1$, whereby the inner condition is Π_1^1 (as it is the negation of $(x, y) \in F$ that occurs upon expanding the implication), and so the entire condition is Π_1^1.

So, the choice set C is Σ_2^1, and total because there exists by the classical result mentioned above a homeomorphism of 2^ω into any non-empty perfect subset of any Polish space (note that this is the only part of the proof where the assumption that F is perfect-valued is used; in particular, we could strengthen the statement to allow $F(x)$ to be in any class of sets that is known to contain a homeomorph of Cantor space. For instance, we can just impose that $F(x)$ be uncountable and have the perfect set property; it follows from sufficiently large cardinals that every projective set has the perfect set property; in ZFC we have the perfect set property for Σ^1_1 sets [13, Theorem 12.2], and with a measurable cardinal we also have it for Σ^1_2 [13, Exercise 27.14]). Thus, with a measurable cardinal we obtain the required universally measurable selection. □

Remark. The above proof is very uneconomical in the computations of the complexities of many of the sets involved, particularly the set H of homeomorphisms in $C(2^\omega)$, which as is shown in [15] is just a G_δ set (and thus, curiously, is itself a Polish space) using countable quantifications over a countable basis of 2^ω in place of our uncountable quantifications over 2^ω itself, though a fair amount of very elementary but convoluted argument is required to show that e.g. the existence of a continuous inverse can be established in this way; the naive approach taken above has its usual advantage of subsuming many such difficulties.

We can also consider a parametrized version of the other important universality property of the Cantor space, which is that every compact metric space is a continuous image of the Cantor space.

In particular, we consider Srivastava's selection theorem for compact-valued multifunctions [22, Theorem 4.2]. We replace the arbitrary measurable space involved there with a standard Borel space, and replace the measurability with respect to its σ-algebra with universal measurability, yielding the following weakened statement.

Theorem 4.2.5. Let X, Y be Polish, and $F : X \rightrightarrows Y$ a compact-valued, Borel-measurable multifunction. Then, there is a map $f : X \times 2^\omega \to Y$, such that:

- For each $x \in X$, $f(x, \cdot)$ is continuous and maps $2^\omega \twoheadrightarrow F(x)$.
- For each $c \in 2^\omega$, $f(\cdot, c)$ is universally measurable.

Proof (with MC*).* We first note that we do not need the explicit tree structure of the Cantor space for this proof, rather the only necessary properties are that 2^ω is compact, and universal for compact metric spaces as mentioned above.

We begin with the same remark as in the original proof: we may assume Y is compact, since any Polish space embeds into a compact metric space (the Hilbert cube). Then, let S be the space of continuous functions $2^\omega \to Y$, with the topology arising from the supremum metric. This is a Polish topology, so S is a Polish space and in light of the first condition on f, we may treat f instead as a map $f : X \to S$, where $x \mapsto f(x, \cdot)$. Write $e : S \times 2^\omega \to Y$ for the map that evaluates $s \in S$ at $c \in 2^\omega$; it is easy to see that this map is continuous. The condition that $f(x, \cdot)$ map into $F(x)$ can be written as

$$\forall (c \in 2^\omega) \Big((x, e(f(x), c)) \in \mathrm{Gr}(F) \Big),$$

and the condition that $f(x,\cdot)$ be onto $F(x)$ could naively be written as

$$\forall(y\in Y)\exists(c\in 2^\omega)\Big((x,y)\in\mathrm{Gr}(F)\to e(f(x),c)=y\Big).$$

However, this is a Π^1_2 condition, which we would like to avoid (as Π^1_2 sets are not easily uniformized). Fortunately, since e is continuous, it suffices for its image to contain a countable dense subset of $F(x)$, which conveniently exists because $F(x)$ is a compact metric space. So, let $\{y_i^{(x)}\}_{i\in\omega}$ be a countable dense subset of $F(x)$, then the condition that $f(x,\cdot)$ be onto $F(x)$ can be better written,

$$\forall(i\in\omega)\exists(c\in 2^\omega)\Big((x,y)\in\mathrm{Gr}(F)\to e(f(x),c)=y_i^{(x)}\Big).$$

We construct a choice set $A\subseteq X\times S$ by declaring $(x,s)\in A$ if and only if

$$\forall(c\in 2^\omega)\Big((x,e(s,c))\in\mathrm{Gr}(F)\Big),$$

and

$$\forall(i\in\omega)\exists(c\in 2^\omega)\Big((x,y)\in\mathrm{Gr}(F)\to e(s,c)=y_i^{(x)}\Big).$$

Since $\mathrm{Gr}(F)$ is Borel and e is continuous, all the inner conditions here are Borel, so the first condition collectively is Π^1_1, and the second is Σ^1_2. Hence, $A\in\Sigma^1_2(X\times S)$. We also want to show that each section of A is non-empty; i.e. there always exists a continuous map $2^\omega\twoheadrightarrow F(x)$ for any x. But this is immediate from the universality of 2^ω mentioned above, since $F(x)$ is compact.

By Kondô's theorem, A then has a Σ^1_2 uniformization $f:X\to S$, which is thus Δ^1_2-measurable as a function. It remains to show that $f(\cdot,c)$ is universally measurable for each c. That is, we want to show that for fixed c, the set of x for which $f(x,c)\in U$ for some open $U\subseteq Y$ is universally measurable. This condition is equivalently $e(f(x),c)\in U$, or $(f(x),c)\in e^{-1}(U)$, the latter set being open since e is continuous. The set of s for which $(s,c)\in e^{-1}(U)$ for fixed c is open since it is a section of an open set, and so the set of x in question is $f^{-1}(V)$ for some open V, and thus is Δ^1_2. Assuming a measurable cardinal, $f(\cdot,c)$ is then universally measurable. □

This sort of theorem can be seen as generalizing the theorems finding a countable family of selections that "fills up" a choice set (either exactly when the sections are countable, or by being dense in each section when they are not), to doing this filling exactly even when the sections are uncountable sets. A useful survey of such theorems, due to Mauldin, can be found in [17].

4.2.4 Stochastic Programming

Here, we consider a parametrized version of a broad class of optimization problems, given by the following model (taken from [7]): let X,U,Y be Banach spaces (later assumed separable as well, which ensures that they are Polish spaces as well), $Q_x \subseteq X$ and $Q_u \subseteq U$, $T : Q_x \times Q_u \rightrightarrows Y$ a multifunction, and $S : Q_x \times Q_u$ an objective function. Then, we are interesting in maximizing $S(x,u)$ subject to $(x,u) \in Q_x \times Q_u$, and $0 \in T(x,u)$ (perhaps a more reasonable-seeming constraint would be to have T be single-valued and just require $T(x,u) = 0$, but this version, as noted by Engl in [7], subsumes that and also covers inequality constraints, for instance $T(x,u) \geq 0$, without any further modification).

The parametrization (the description taken from the same source) is over an arbitrary measurable space $(\Omega, \mathcal{F})$, where $Q_x : \Omega \rightrightarrows X$ and $Q_u : \Omega \rightrightarrows U$ are now measurable and closed-valued multifunctions. We define the multifunction $Q_x \times Q_u : \Omega \rightrightarrows X \times U$ by $(Q_x \times Q_u)(\omega) = Q_x(\omega) \times Q_u(\omega)$, and the topological space $\mathrm{CB}(Y)$ as the space of closed, bounded subsets of Y equipped with the Hausdorff distance.

Then, we take the constraint function to be $T : \mathrm{Gr}(Q_x \times Q_u) \to \mathrm{CB}(Y)$, "measurable" in the sense that for all $(x,u) \in X \times U$ and $D \subseteq Y$ open, we have that the set

$$\{\omega \in \Omega : x \in Q_x(\omega), u \in Q_u(\omega), T(\omega, x, u) \cap D \neq \emptyset\}$$

is in $\mathcal{F}$ (i.e. is measurable in Ω). Similarly, we take the objective function to be $S : \mathrm{Gr}(Q_x \times Q_u) \to \mathbb{R}$, likewise "measurable" in the sense that for $(x,u) \in X \times U$ and $D \subseteq \mathbb{R}$ open, the set

$$\{\omega \in \Omega : x \in Q_x(\omega), u \in Q_u(\omega), S(\omega, x, u) \in D\}$$

is in $\mathcal{F}$. The functions T and S are also presumed jointly continuous in x and u, for a fixed $\omega \in \Omega$ in the first argument.

Let us write $P(\omega, x, u)$ for the proposition that is the conjunction of $x \in Q_x(\omega)$, $u \in Q_u(\omega)$, and $0 \in T(\omega, x, u)$ (i.e. the criterion for (x,u) to be an "admissible" point for the maximization problem at outcome ω). We are then interested in finding measurable functions $\overline{x} : \Omega \to X$ and $\overline{u} : \Omega \to U$ such that for all $\omega \in \Omega$, we have $P(\omega, \overline{x}(\omega), \overline{u}(\omega))$, and

$$S(\omega, \overline{x}(\omega), \overline{u}(\omega)) = \sup_{x,u \text{ s.t. } P(\omega,x,u)} S(\omega, x, u).$$

(In [7], part of the desired result is also that this latter supremum function be measurable as a function $\Omega \to \mathbb{R}$, but this is a property of the described problem, not the solution, and as shown there under the stated assumptions this function is always measurable anyway, so we do not examine this part further.) Such a pair of functions is termed a *stochastic solution* to the maximization problem framed above. A *deterministic solution* at a particular random outcome $\omega \in \Omega$ is just a pair (x,u) that at this ω satisfies $P(\omega, x, u)$ and the above maximization condition (i.e. a solution to the non-randomized problem, as described briefly earlier, induced by fixing ω). The main question here is then whether the existence of a deterministic solution at each ω ensures the existence of a stochastic solution (which must be measurable over Ω

and thus cannot just consist of deterministic solutions “glued together” arbitrarily). The below theorem from [7] answers this question.

Theorem 4.2.6. In the arrangement described above, the maximization problem defined by Q_x, Q_u, S, T has a stochastic solution if and only if it has a deterministic solution at each $\omega \in \Omega$ (a stochastic solution immediately yields deterministic solutions at each ω, so only the “if” implication is non-trivial).

To cast this in the terms we have been dealing with previously, we take Ω to be a Polish space and $\mathcal{F}$ to be its Borel sets (so that Ω as a measurable space is a standard Borel space), and note that U, X, Y are all Polish spaces already (being separable Banach spaces). Thus, $Q_x \times Q_u$ is a Borel measurable multifunction, and so its graph $G = \mathrm{Gr}(Q_x \times Q_u) \subseteq \Omega \times X \times U$ is also Borel. As usual, we will prove the result giving universally measurable functions $\overline{x}$, $\overline{u}$, not Borel measurable ones.

Proof (with MC*).* We construct both functions at once as a single function $(\overline{x}, \overline{u}) : \Omega \to X \times U$ (notice that though taking X and U to be separate spaces is important in the original formulation, we are basically just interested in the Polish space $X \times U$). If we produce a universally measurable such function, then its components will give the desired universally measurable $\overline{x}$ and $\overline{u}$. Our choice set $C \subseteq \Omega \times X \times U$ then consists of $(\omega, x, u) \in \Omega \times X \times U$ that satisfy $P(\omega, x, u)$ and also the maximization condition:

$$S(\omega, x, u) = \sup_{x', u' \text{ s.t. } P(\omega, x, u)} S(\omega, x', u').$$

We can rewrite this as

$$\forall\, ((x', u') \in X \times U) \Big(S(\omega, x', u') \leq S(\omega, x, u) \Big),$$

where since S is continuous in the last two arguments jointly, the inner condition is a Borel condition, and so this entire condition is Π^1_1.

It remains to establish the complexity of $P(\omega, x, u)$, which is the condition that $x \in Q_x(\omega)$, $u \in Q_u(\omega)$, and $0 \in T(\omega, x, u)$. Since Q_x and Q_u are Borel maps, they have Borel graphs, and so the conditions $x \in Q_x(\omega)$ and $u \in Q_u(\omega)$ are Borel. Since T as a multifunction $\Omega \rightrightarrows X \times U \times Y$ is Borel measurable, it has a Borel graph. The set of (ω, x, u) that satisfy P is the projection of the intersection of this graph with the subset of $\Omega \times X \times U \times Y$ with 0 in the last coordinate (in the general Polish space context, we can view 0 as an arbitrary distinguished point). So, the set of points satisfying $P(\omega, x, u)$ is Σ^1_1. Thus, the choice set C is Δ^1_2 (the intersection of a Σ^1_1 and a Π^1_1). We have that C as a multifunction $\Omega \rightrightarrows X \times U$ is total by assumption (according to the existence of deterministic solutions), and so with a measurable cardinal we have the desired universally measurable selection. □

4.3 Inverse and Implicit Function Theorems

We saw at the beginning that the question of the existence of left inverses for functions was basically identical to the question of selection from the graphs of those functions. In particular, our work earlier immediately implies that any Borel map has a universally measurable left inverse. We now consider a few variations on this theme.

4.3.1 Jankoff's Right Inverse Theorem

The most natural place to begin is, of course, with right inverses.

Theorem 4.3.1 (Jankoff). Let X, Y be Polish spaces, and $F : X \to Y$ be a Borel map that is surjective onto $E \in \Sigma^1_1(Y)$ (in the original version, X is a Suslin space, or equivalently an analytic subspace of some Polish space, but any such space is a continuous image of a Polish space, and composing F with this continuous map puts us in the above setting). Then, there is a map $G : Y \to X$ such that $F \circ G$ is the identity on E, G is $\sigma^1_1(Y)$-measurable, and $G(E) \in \sigma^1_1(X)$.

We will give a proof of this theorem with the measurability conclusion weakened to universally measurability in the conclusion, and likewise the condition on $G(E)$ weakened to universal measurability, assuming the existence of a measurable cardinal.

Proof (with MC*).* Let us construct a subset $A \subseteq Y \times X$ by its sections, $A(y) = F^{-1}(\{y\})$ if this set is non-empty, and $A(y) = X$ otherwise (in other words, this is the graph of the inverse multifunction of F, with empty "slices" filled in). Then, every section over $y \in Y$ of this graph is non-empty by construction. Note that $A = \mathrm{Gr}(F) \cup E^c \times X$, where $\mathrm{Gr}(F)$ is interpreted to be the actual graph of F but flipped to lie in $Y \times X$ rather than $X \times Y$.

Since F is a Borel map, $\mathrm{Gr}(F)$ is Borel, and $E^c \times X$ is Π^1_1, thus $A \in \Pi^1_1$. Then, $\mathsf{UnivSel}(\Pi^1_1)$ gives a universally measurable selection of A, and such a selection G clearly satisfies the condition of being a right inverse for F (with no measurable cardinal assumption, we would get a Δ^1_2-measurable selection, which is also weaker than the result in the theorem, as σ^1_1 is strictly contained in Δ^1_2).

Also, recall that we produce this selection by starting with a uniformization from the hypothesis $\mathsf{Unif}(\Pi^1_1)$, and thus $\mathrm{Gr}(G) \in \Pi^1_1$. We have $G(E) = \pi_X(\mathrm{Gr}(G) \cap (E \times X))$, where $E \times X$ is Σ^1_1, so this belongs to $\exists^{\omega^\omega} \Delta^1_2 = \Sigma^1_2$ and so is universally measurable assuming a measurable cardinal.

Note that this method still works under weaker assumptions; in particular, we do not need F to be a Borel map, but only to have its graph be in Σ^1_2 (and assuming Woodin cardinals allows this bound on complexity to be pushed further up the projective hierarchy, and likewise the bound on the complexity of E). □

The significantly more involved standard proof can be found in [3]. In the final remark of the proof we see another feature that is common to measurable selection proofs conducted with the large cardinal approach: it is possible to incrementally weaken the assupmtions of our results (moving some measurability assumption up the σ^1_α sequence of σ-algebras) by assuming more large cardinals (the details of this are usually sufficiently clear that we may omit them later).

4.3.2 Filippov's Implicit Function Theorem

We now turn to a selection theorem that (under various permutations of differing assumptions) generalizes the above two results, and is central to the theory of optimal control. Below we work with the version in [24, Theorem 2.3.13], which generalizes Filippov's original result slightly (and casts it into a setting that is convenient to express in terms of Polish spaces), but there have been a variety of generalizations and variations in the literature. From the version in [24], we replace real spaces $\mathbb{R}^k$ with arbitrary Polish spaces to obtain the following.

Theorem 4.3.2 (Generalized Filippov Selection Theorem). Let X, Y, Z be Polish, and let $F : X \rightrightarrows Y$ be a total multifunction with $\mathrm{Gr}(F) \in \mathcal{L}(X) \otimes \mathcal{B}(Y)$, and let $g : X \times Y \to Z$ be a function that is $\mathcal{L}(X) \otimes \mathcal{B}(Y)$-measurable. Then, for any Lebesgue-measurable map $v : X \to Z$, if we have

$$v(x) \in \{g(x,y) : y \in F(x)\}$$

for almost all x (in the Lebesgue sense), then there is a Lebesgue-measurable function $u : X \to Y$ that is a selection for F almost everywhere, and we also have $g(x, u(x)) = v(x)$ for almost all x.

We must adapt this somewhat to be accessible by our methods. First, we remove all Lebesgue almost-all conditions, then convert all references to the Lebesgue σ-algebra $\mathcal{L}$ to the analytic σ-algebra σ^1_1—this is a stronger assumption (σ^1_1 sets are Lebesgue measurable). Lastly, we change the conclusion to universal measurability (this is even slightly stronger than the reference version above).

Theorem 4.3.3 (Adapted Filippov Selection Theorem). Let X, Y, Z be Polish, and let $F : X \rightrightarrows Y$ be a total multifunction with $\mathrm{Gr}(F) \in \sigma^1_1(X \times Y)$, and let $g : X \times Y \to Z$ be a function that is σ^1_1-measurable. Then, for any σ^1_1-measurable map $v : X \to Z$, if we have

$$v(x) \in \{g(x,y) : y \in F(x)\}$$

for all x, then there is a universally measurable function $u : X \to Y$ that is a selection for F, and such that $g(x, u(x)) = v(x)$ for all x.

Proof (with MC*).* An equivalent condition is that u be a universally measurable selection on the set

$$A = \{(x, y) : g(x, y) = v(x)\} \cap \mathrm{Gr}(F),$$

where by assumption each section $A(x)$ of A is non-empty. We have a direct assumption on the complexity of $\mathrm{Gr}(F)$, and we can write the first set above as

$$\{(x, y) : g(x, y) = v(x)\} = \pi_{X\times Y}(\mathrm{Gr}(g) \cap (\mathrm{Gr}(v) \times Y)),$$

where $\mathrm{Gr}(v) \times Y$ takes the Cartesian product in the middle coordinate, so that $\mathrm{Gr}(g)$ and $\mathrm{Gr}(v) \times Y$ can both be taken to lie in $X \times Y \times Z$. Since all graphs in question are σ_1^1, we have that A is σ_1^1 as well, and thus has a universally measurable selection, assuming a measurable cardinal. □

Indeed, for use with our proof, this formulation is somewhat stilted, because we are actually primarily interested in the graphs of F, g, and v, not their measurability. In particular, it suffices to assume $\mathrm{Gr}(F), \mathrm{Gr}(g), \mathrm{Gr}(v) \in \Sigma_2^1$ for this proof to go through.

4.4 Approximate Optimization Theorems

In this section we consider selection theorems where the selections must satisfy an explicit metric property, an approximation property of an objective function depending on an error parameter ϵ.

4.4.1 Uniformly Approximating Objective Function Extrema

Definition 4.4.1. For X, Y Polish spaces, $D \subseteq X \times Y$, and $f : D \to [-\infty, \infty]$ a function, define for $x \in \pi_X(D)$:

$$f_*(x) = \inf_{y\in D_x} f(x, y),$$
$$f^*(x) = \sup_{y\in D_x} f(x, y).$$

For our purposes, the distinction here is not at all important (i.e. we can negate f to switch from infimum to supremum without affecting any descriptive complexity considerations); we will thus write our theorems in terms of one of these functions to conform with the source material, but analogous statements will always hold for the other function as well. We will be looking below at these functions when f is a Borel map. It is easy to see that f_* and f^* need not be Borel when f is Borel.

Definition 4.4.2. For X a Polish space and $D \subseteq X$, a function $f : D \to [-\infty, \infty]$ is *lower semianalytic* if for all $c \in \mathbb{R}$ and $c = \pm\infty$, any of the following equivalent sets of conditions hold:

- $\{x \in X \mid f(x) < c\} \in \Sigma_1^1(X)$.

- $\{x \in X \mid f(x) \leq c\} \in \Sigma^1_1(X)$.
- $\{x \in X \mid f(x) > c\} \in \Pi^1_1(X)$.
- $\{x \in X \mid f(x) \geq c\} \in \Pi^1_1(X)$.

It is *upper semianalytic* if any of these hold with the inequalities reversed (or with Σ^1_1 and Π^1_1 swapped).

In particular, we note that any function of either class is automatically σ^1_1-measurable. We have

$$f^*(x) > t \leftrightarrow (\exists y \in Y)(f(x,y) > t),$$

hence if f is Borel, then f^* is upper semianalytic, and by the same token f_* is lower semianalytic. And so, both of these functions are σ^1_1-measurable (and, as will be useful in relation to a further variant of this theorem, if f is lower semianalytic then f_* is still lower semianalytic, and likewise for upper semianalyticity and f^*).

We investigate here the question of approximating the points realizing the values of the extremal functions f_* and f^* on a planar set by a uniformly converging sequence of regular functions, as in the below theorem, considered in [1].

Theorem 4.4.3. Let X, Y be Polish spaces, $G : X \rightrightarrows Y$ with $\mathrm{Gr}(G) \in \Sigma^1_1(X \times Y)$. Let $u : X \times Y \to \mathbb{R}$ be a non-negative Borel function. Then, there is a uniformly convergent sequence $\phi_i : X \to Y$ of maps that are universally measurable such that $\lim_{i\to\infty} u(x, \phi_i(x)) = u^*(x)$, where this limit holds uniformly in x over those x for which $u^*(x) < \infty$.

The same statement holds symmetrically for u_* as well.

Proof (with MC*).* As usual for countable sequences of selections, we construct a single function $\Phi : X \to Y^\omega$ and use its coordinates as our ϕ_i; if Φ is universally measurable then so are all of the ϕ_i. The condition

$$\lim_{i\to\infty} u(x, \phi_i(x)) = u^*(x) = \sup_{y \in G(x)} u(x,y)$$

can be written

$$\forall (y \in G(x)) \Big(u(x,y) \leq \lim_{i\to\infty} u(x, \phi_i(x)) \Big).$$

We define the function $\pi : Y^\omega \times \omega \to Y$ where $\pi(r,i)$ is the projection of r onto the ith coordinate, so that we can write the inner condition as

$$\forall (m \in \mathbb{N}_+) \exists (n \in \mathbb{N}) \forall (i > n) \Big(u(x,y) < u(x, \pi(\Phi(x), i)) + \frac{1}{m} \Big),$$

so the graph of Φ is a subset of the set (x,z) such that

$$\forall (y \in G(x)) \forall (m \in \mathbb{N}_+) \exists (n \in \mathbb{N}) \forall (i > \mathbb{N}) \Big(u(x,y) < u(x, \pi(z,i)) + \frac{1}{m} \Big),$$

let us call this set A. Since u is a Borel function and π is continuous, this entire condition up to the quantifier $\forall(y \in G(x))$ is Borel (in the collection of all variables present). And, since $\mathrm{Gr}(G) \in \Sigma^1_1$, we have that A is Π^1_1.

So, if Φ is a selection of A, then $\lim_{i\to\infty} u(x, \phi_i(x)) = u^*(x)$ for all x. Now, we want to ensure that over the x for which $u^*(x) < \infty$, this convergence is uniform. Note that general uniform convergence is not a condition that is easily handled by the sorts of constructions we have been doing, since it cannot be written as the conjunction of individual local conditions at each x. However, in this case we can use a simple restriction to overcome this. Let us choose a sequence of arbitrary functions α_i such that $u(x, \alpha_i(x)) \to u^*(x)$ uniformly when $u^*(x) < \infty$ (for each x, there is a sequence of a_i so that $u(x, a_i) \to u^*(x)$, and we can just form the α_i by choosing the a_i from these sequences for each x so as to make the convergence uniform). This means that for any $k \in \omega$, there is $N(k) \in \omega$ such that whenever $i > N(k)$, we have $u(x, \alpha_i(x)) > u^*(x) - \frac{1}{k}$ for all x.

Then, we simply impose this specific "regime" of uniform convergence on the graph of Φ. Let I be the set where $u^*(x) < \infty$; since u^* is upper semianalytic, we have $I \in \Sigma^1_1(X)$. Now, we form another choice set $B \subseteq X \times Y^\omega$ where now $(x, z) \in B$ if and only if

$$\forall(k \in \omega)\forall(i \in \omega)\left(x \in I \text{ and } i > N(k) \to u(x, \pi(z, i)) > u^*(x) - \frac{1}{k}\right),$$

which is a Π^1_1 condition.

We will have the desired result if Φ is a selection from $A \cap B$, which is a Π^1_1 set and is total since the graph of the function $(\alpha_1, \alpha_2, \dots)$ chosen arbitrarily earlier is in this set. With a measurable cardinal, we obtain the required universally measurable selection. □

4.4.2 The Exact Selection Theorem

It is also possible to provide a more explicit construction of such a sequence of measurable functions uniformly approximating the extrema of f. We consider here a selection theorem central to the area of dynamic programming, which concerns selections closely approximating the infimum of a function over sections of a set. The following is the basic form of the exact selection theorem, first proved in [4].

Theorem 4.4.4 (Borel Exact Selection Theorem). Let X, Y be Polish spaces, $D \subseteq X \times Y$ a Borel subset, $f : D \to [-\infty, \infty]$ a Borel measurable function, and define the sets

$$G = \pi_X(D),$$
$$I = \{x \in G \mid \exists(y \in D_x)(f(x, y) = f_*(x))\}.$$

Then, both of these sets are universally measurable. Moreover, for each $\epsilon > 0$, there exists a universally measurable selection $\varphi : X \to Y$ of D such that:

- If $x \in I$, then $f(x, \varphi(x)) = f_*(x)$.
- If $x \in G \setminus I$ and $f_*(x) > -\infty$, then $f(x, \varphi(x)) \leq f_*(x) + \epsilon$.
- If $x \in G \setminus I$ and $f_*(x) = -\infty$, then $f(x, \varphi(x)) \leq -\frac{1}{\epsilon}$.

This formulation generalizes the contexts of a fairly wide range of other selection theorems, perhaps the most prominent among them being the Dubins-Savage selection theorem from dynamic programming (a good reference and proof can be found in [14]), where the function f is assumed bounded, Y compact, and D of the form $B \times Y$ for some $B \subseteq X$ Borel, so that $G = I = B$ and the infimum is achieved on every section, so ϵ is irrelevant.

Note also that the specific function $-\frac{1}{\epsilon}$ is not essential here, any continuous function decreasing to $-\infty$ as $\epsilon \to 0$ from above will do. Indeed, we will see in the proof below that much of the structure of this problem is not strictly essential.

Proof (with MC*).* We immediately have that G is analytic and thus universally measurable, and the condition $f(x, y) = f_*(x)$ used in defining I can be written as

$$\Big(f(x, y) = f_*(x)\Big) \leftrightarrow (\forall z \in Y)\Big((x, z) \in D \to f(x, y) \leq f(x, z)\Big)$$

The inner condition is Borel (as D is Borel and f is a Borel map), so $f(x, y) = f_*(x)$ is a Π^1_1 condition. The set I is the projection of the (x, y) for which $f(x, y) = f_*(x)$ onto X, and so $I \in \Sigma^1_2(X)$, and so is also universally measurable.

To construct the measurable selection, our earlier example of handling piecewise functions will be useful, as the desired properties of the selection are phrased as independent conditions over disjoint possible values of x. We define a separate choice set for each condition above:

$$\begin{aligned} C_1 &= \{(x, y) : f(x, y) = f_*(x)\}, \\ C_2 &= \{(x, y) : f(x, y) \leq f_*(x) + \epsilon\}, \\ C_3 &= \left\{(x, y) : f(x, y) \leq -\frac{1}{\epsilon}\right\}. \end{aligned}$$

The set C_1 is the set of points satisfying the condition $f(x, y) = f_*(x)$, which we saw above was Π^1_1. The set C_2 is the set of points satisfying the condition $f(x, y) \leq f_*(x) + \epsilon$, which can be written as

$$\forall(n \in \omega_+)\exists(z \in Y)\Big((x, z) \in D \text{ and } f(x, y) < f(x, z) + \epsilon + \frac{1}{n}\Big),$$

where the inner condition is Borel and so C_2 is Σ^1_1. Lastly, the set C_3 is the set of points satisfying $f(x, y) \leq -\frac{1}{\epsilon}$, which is Borel since f is a Borel map. In particular, all C_i are Σ^1_2, so from Kondô's theorem in conjunction with 3.1.3 we have that each C_i has a Δ^1_2-measurable selection, call it ϕ_i for $i = 1, 2, 3$.

Now, define the domain sets

$$
\begin{aligned}
D_1 &= I, \\
D_2 &= (G \setminus I) \cap \{x : f_*(x) > -\infty\}, \\
D_3 &= (G \setminus I) \cap \{x : f_*(x) = -\infty\}.
\end{aligned}
$$

Since f_* is lower semianalytic, the condition $f_*(x) > -\infty$ is Π^1_1, and the condition $f_*(x) = -\infty$ is the conjuction of $f_*(x) \leq -\infty$ and $f_*(x) \geq -\infty$, which are Σ^1_1 and Π^1_1 respectively, so $f_*(x) = -\infty$ is Δ^1_2. We also have $I \in \Sigma^1_2$ and $G \in \Sigma^1_1$, and so $I, G \setminus I \in \sigma^1_2$. In particular, all the sets involved in defining D_1, D_2, and D_3 are σ^1_2, thus D_1, D_2, and D_3 are themselves σ^1_2.

Now, we construct the selection ϕ by defining it piecewise, taking $\phi(x) = f_i(x)$ whenever $x \in D_i$, and completing the selection by an arbitrary σ^1_2-measurable function on $X \setminus G$ (equivalently $G \setminus (D_1 \cup D_2 \cup D_3)$). Then, the preimage of any Borel set under ϕ is σ^1_2 in X, since it is equal to the union of the preimages under the f_i and the arbitrary function above, all of which are σ^1_2-measurable. Thus, ϕ is σ^1_2-measurable, as is its restriction to G, since G is Σ^1_1. Assuming a measurable cardinal, we then have that ϕ is universally measurable, as desired. □

A few extensions are possible without too much difficulty. First, we can weaken the restriction on f to being lower semianalytic, since then f_* is still lower semianalytic and f is σ^1_1-measurable, so the set I is σ^1_1-measurable and still universally measurable. In the selection construction, the C_i are all still at most Σ^1_2, and thus the proof still goes through. Also, it is easy to check that it is enough to assume $D \in \Sigma^1_1$ (G is still Σ^1_1, and the complexity of C_1 and thus I does not change). This strengthened version is derived by classical means in [2], where it is the main ingredient in the proof of a parametrized/randomized version of Fan's minimax theorem from continuous game theory.

References

[1] Aliprantis, C. D., and Border, K. *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 3rd ed. Springer, 2007.

[2] Bertsekas, D. P., and Shreve, S. E. Alternative theoretical frameworks for finite horizon discrete-time stochastic optimal control. *SIAM Journal on Control and Optimization 16*, 6 (November 1978), 953–977.

[3] Bogachev, V. I. *Measure Theory: Volume I*. Springer, 2007.

[4] Brown, L., and Purves, R. Measurable selections of extrema. *Annals of Statistics 1*, 5 (1973), 902–912.

[5] Burgess, J. P. From preference to utility: A problem of descriptive set theory. *Notre Dame Journal of Formal Logic 26*, 2 (April 1985), 106–114.

[6] Castaing, C., and Valadier, M. *Convex Analysis and Measurable Multifunctions*, vol. 580 of *Lecture Notes in Mathematics*. Springer, 1977.

[7] Engl, H. W. Existence of measurable optima in stochastic nonlinear programming and control. *Journal of Applied Mathematics and Optimization 5* (1979), 271–281.

[8] F ollmer, H., and Schied, A. *Stochastic Finance: An Introduction in Discrete Time*, 2nd ed. No. 27 in de Gruyter Studies in Mathematics. Walter de Gruyter, 2004.

[9] Friedman, H. Does normal mathematics need new axioms? Online Lecture Notes, May 2000.

[10] Friedman, H. What you cannot prove: Before 2000. Online Lecture Notes, December 2005.

[11] Jackson, S., and Mauldin, R. D. Borel measurable selections of paretian utility functions. *Journal of Mathematical Economics 23* (1994), 361–378.

[12] Jech, T. *Set Theory*, 3rd ed. Springer, 2006.

[13] Kanamori, A. *The Higher Infinite: Large Cardinals in Set Theory from Their Beginnings*, 2nd ed. Springer, 2003.

[14] Maitra, A. Discounted dynamic programming on compact metric spaces. *Sankhyā: The Indian Journal of Statistics, Series A 30*, 2 (1968), 211–216.

[15] Maitra, A., Rao, B. V., and Srivatsa, V. V. Some applications of selection theorems to parametrization problems. *Proceedings of the American Mathematical Society 104*, 1 (September 1988), 96–100.

[16] Martin, D. A., and Solovay, R. M. Internal cohen extensions. *Annals of Mathematical Logic 2*, 2 (1970), 143–178.

[17] Mauldin, R. D. Some selection theorems and problems. In *Proceedings of Measure Theory Oberwolfach 1979* (July 1979), D. K ozlow, Ed., Springer, pp. 160–165.

[18] Mauldin, R. D. Measurable representations of preference orders. *Transactions of the American Mathematical Society 275*, 2 (February 1983), 761–769.

[19] Moschovakis, Y. N. *Descriptive Set Theory*, 2nd ed. American Mathematical Society, 2009.

[20] Normann, D., and Fenstad, J. On absolutely measurable sets. *Fundamenta Mathematicae 81*, 2 (1974), 91–98.

[21] Shelah, S., and Woodin, W. H. Large cardinals imply that every reasonably definable set of reals is lebesgue measurable. *Israel Journal of Mathematics 70*, 3 (1990), 381–394.

[22] Srivastava, S. M. Selection and representation theorems for σ-compact valued multifunctions. *Proceedings of the American Mathematical Society 83*, 4 (December 1981), 775–780.

[23] Srivastava, S. M. *A Course on Borel Sets.* Springer, 1998.

[24] Vinter, R. *Optimal Control.* Modern Birkh auser Classics. Birkh auser, 2000.

[25] von Neumann, J. On rings of operators: reduction theory. *Annals of Mathematics 50*, 2 (April 1949), 401–485.

[26] Wagner, D. H. Survey of measurable selection theorems. *SIAM Journal of Control and Optimization 15*, 5 (1977), 176–219.

[27] Wesley, E. Borel preference orders in markets with a continuum of traders. *Journal of Mathematical Economics 3*, 2 (1976), 155–165.