Obtaining Woodin’s Cardinals

P. D. Welch

Abstract. Since the 1980’s work on Projective Determinacy and AD\(_L^{\mathbb{R}}\) the concept of Woodin cardinal has become to be seen as central in the theory of large cardinals and inner model theory. The use by Woodin himself of a background assumption in many arguments that the universe contains unboundedly many such cardinals again draws attention to the centrality of this concept.

As is well known the Reflection Principles dating to a more classical era only provide large cardinals consistent with \(V = L\), and not the wherewithal for such theorems on absoluteness under set forcing that Woodin has proven.

We discuss here a reflection principle derived from weak sub-compactness that implies the existence of a proper class of measurable Woodin cardinals - thus providing adequate background assumptions for many of Woodin’s absoluteness results in his work.

O, there has been much throwing about of brains.

Guildenstern; Hamlet II.2

Contents

1. Introduction 2
2. The Cantorian versus the Zermelian realms 2
3. Reflection 4
4. The Ackermannian realm and reflection 5
5. Global Reflection 7
6. Is GRP a reflection principle? 10
7. Strengthening GRP? 13
8. Conclusion 15
References 15

2010 Mathematics Subject Classification. Primary 03E55, Secondary 03E05.
Key words and phrases. Woodin cardinals, reflection.

©0000 (copyright holder)
1. Introduction

This article is not a history of the origins of Woodin’s notion of large cardinal now named after him, which was so central in the arguments used to establish Projective Determinacy by Martin and Steel, and Woodin himself for $\text{AD}^L(R)$; this history is told elsewhere - see for example [9]. The ubiquity of Woodin cardinals is attested by the literature today not on just determinacy issues, but on so very many of the consistency results with which we gauge the strength of set theoretic objects. However it is not due to the power of infinitely many Woodin cardinals to settle the question of definable determinacy (whether taken in the PD or $\text{AD}^L(R)$ form) or as Woodin has written [26] to give us as good as complete a theory of the hereditarily countable sets, HC, as possible, but rather the use of unboundedly many Woodin cardinals throughout the ordinals as an, again uniquitous, background assumption for many of his, and others’ work on, for example, establishing the absoluteness of many properties of our universe, most typically into imagined generic extensions of the universe $V$ by set-sized forcing notions.

If a hoped for “reduction in incompleteness” over our standard axioms of ZFC is to be achieved by the adoption of new axioms, and if we are to attempt to fully justify those axioms, then arguing for an axiom that yields a proper class of Woodin cardinals is an excellent place to start. Let UW abbreviate this axiom (for ‘there exist Unboundedly many Woodin cardinals’).

What conception of set or universe of sets can we have that will deliver this for us? That ZFC could and should be extended was famously pleaded for by Gödel in [7] which is by now a locus classicus:

“the axioms of set theory by no means form a system closed in itself, but, quite on the contrary, the very concept of set on which they are based suggests their extension by new axioms which assert the existence of still further iterations of the operation ‘set of’.”

“...[The ZFC axioms may be] supplemented without arbitrariness by new axioms which only unfold the content of the concept of set”.

A discussion of the nature of intrinsic necessity in the words of Gödel, or of intrinsic versus extrinsic justification should probably now be made but I shall short-circuit this by referring the reader to Koellner’s article [10] and discussion. It is not my intention to wade in here. The discussion here is about what possible “conception of set” could lead to UW.

2. The Cantorian versus the Zermelian realms

Cantor’s discoveries and advances were made as a mathematician would work; in a non-formalised manner (and even that phrase is anachronistic). His viewpoint concerning the world of order types and cardinalities would be formed in an intuitive fashion. In the past it was stated that Cantor’s views were that of a ‘naive set theorist’, a description not as usually used, but with rather too much emphasis on ‘naive’. Now, however we realise that in fact he was quite aware of the pitfalls of what we would call the set/class distinction. At different stages of his career he

---

1This paper was partially prepared whilst the author was a Simons Foundation Visiting Fellow at the Isaac Newton Institute for Mathematical Sciences in the programme ‘Mathematical, Foundational and Computational Aspects of the Higher Infinite’ (HIF), supported by EPSRC grant EP/K032208/1, to both of which the author would like to express his gratitude.
used the phrase “The Absolute Infinite” or around the time of the publication of Burali-Forti (1897) - “inconsistent multiplicities”, or later - both.

In a letter to Dedekind (1899)\[5\]:

A multiplicity can be of such a nature, that the assumption of the ‘togetherness’ (‘Zusammenseins’) of a multiplicity’s elements leads to a contradiction, so that it is impossible to conceive the multiplicity as a unity, as a ‘finished thing’. Such multiplicities I call absolute infinites or inconsistent multiplicities.

In the passage here he is aiming for the ‘finished things’ or consistent things, to be sets. The multiplicity of all alephs for example, cannot constitute a finished set and so cannot be assigned a cardinal number. (See [8] for a discussion of this notion of Zusammensein.)

In latter day jargon we should call such things ‘proper classes’. It does not show that Cantor no longer thought of the set theoretic universe as an ‘inaugmentable totality’, which he had called it earlier. (Another description was that it was an ‘absolute maximum’ ([4], pp.410-411)). We cannot be entirely clear what Cantor had in mind when discussing this universe of sets in this pre-formalised manner, but it was clearly different from the view Zermelo was to come to have.\[^2\]

Zermelo’s maturest picture has come down to us from his [27]. The view is that there are only sets and that these satisfy, let us say here, first order ZFC (although Zermelo was concerned to promote a second order view and eschewed the first order formulations of Skolem and others). For Zermelo the only collections are sets. For Zermelo when we do set theory our quantifiers range over a domain of discourse $D$ say. The ‘paradoxes’ show us that the collection $D$ cannot be a member of itself. Hence we can enlarge this domain to a larger domain of discourse $D'$ in which $D$ is a set. Hence we have a never ending sequence of, in his words, ‘normal domains’ which are models of second order ZF, (and hence their ordinal heights are strongly inaccessible cardinals); the sequence of these domains can be indexed by Cantor’s ordinal numbers. Zermelo talks of ‘creative advancement’ as one proceeds through these domains; and that we should talk about such a sequence of domains in some meta-theory. However this meta-theory is never laid out, much beyond the indication that the normal domains should be in a (1-1) correspondence with the Cantorian ordinals. (Readers of [27] will see that I am deliberately ignoring one, inessential for this discussion, aspect of Zermelo’s picture: the presence of a possible non-empty collection of urelements over which a normal domain may also be built. Zermelo’s normal domains encompassed models of $ZF^2$ with or without urelements, but for this paper it is inessential: we shall consider only normal domains without urelements. One should note that also under Zermelo’s second order conception Choice is a logical, as opposed to a set theoretical, axiom. Hence $ZF^2$ is to be regarded as including Choice.)

However these different views give us at least two broad-brush pictures of the universe of sets: a ‘potentialist’ view - Zermelo’s creative advancement, and an

\[^2\]Interestingly from the same letter, recalling that ‘equivalent’ meant bijective, one has a prefiguration of Replacement: “Two equivalent multiplicities are either both “sets” or are both inconsistent”.
`actualist’ view - that the universe of sets is an absolute maximum and an inaug-
mentable totality. It is possible to discuss these views unlinked to any kind of
position concerning platonism or realism.

3. Reflection

A potentialist view makes it hard to appeal to most kinds of reflection principle.
Zermelo cannot consider the whole universe, and reflect on that, since for him
there is always the potential to make the universe yet larger. All that can be
said here is that there are unboundedly many normal domains (which following
Mirimanoff/von Neumann we should now recognise as $V_\alpha$’s) and so a proper class of
strongly inaccessible cardinals. But even this can not be obtained by an adherent
of the Zermelian school as a result of reflection: the statement \( \forall \alpha \in \text{On} \exists \beta > \alpha (ZF^2)^{V_\beta} \) expresses quantification over all ordinals, and this is just what cannot
be done in this organic view. The statement can however be readily obtained by
allowing second order reflection of the whole (‘actual’) universe ($V, \in$).

By allowing domains of all sets and all classes - as formalised by NBG say - we
now are able to second-order quantify over all such classes and formulate reflection
principles that yield Mahlo cardinals, and second order indescribability. The story
of this is familiar enough, that we don’t repeat this here.

The point remains that all such principles only derive cardinals consistent with
$V = L$ - so we may call them intra-constructible.

Gödel stated that he thought all large cardinals could be obtained through
reflection:

The Universe of sets cannot be uniquely characterized (i.e. dis-
tinguished from all its initial segments) by any internal structural
property of the membership relation in it, which is expressible in
any logic of finite or transfinite type, including infinitary logics
of any cardinal number.

(Wang - [24])

Gödel again:

All the principles for setting up the axioms of set theory should
be reducible to a form of Ackermann’s principle: The Absolute
is unknowable. The strength of the principle increases as we get
stronger and stronger systems of set theory. The other principles
are only heuristic principles. Hence the central principle is the
reflection principle, which presumably will be understood better
as our experience increases. (Wang - ibid.)

This seems rather sweeping and our experience with Reflection Principles seems
to go against it. Both on the one hand because Reinhardt specifically noted that
third order reflection with parameters is inconsistent, and on the other, because our
reflection principles have remain stuck in the intra-constructible.

There have been specific attempts to get around this obstacle by restricting the
syntax: Marshall [15] obtains higher order reflection (and large cardinals) by this
method. Tait [21] uses Relativized Cantorian Principles based on certain Existence
Conditions. As motivating conditions these allow him to define certain syntactically
characterised higher order classes of formulae $\Gamma^{(m)}_n$ for $m \geq 2$ (the superscript
indicates that higher order universal quantification must be of at most order \( m \). Tait shows that for \( m = 2 \), \( V_\kappa \) satisfying \( \Gamma_2^{(m)} \) reflection implies that \( \kappa \) is \( n \)-ineffable (in the definition of Baumgartner), and that measurability of \( \kappa \) sufficed to show that \( V_\kappa \) satisfied \( \Gamma_\kappa^{(m)} \)-reflection for all \( n \). This left open the question of whether such principles were extra-constructible. Koellner answered this negatively by showing that if \( \kappa \) is \( \kappa(\omega) \)-Erdös cardinal then for every \( n \) \( V_\kappa \) satisfies \( \Gamma_n^{(m)} \)-reflection. He further answers negatively what was asked, and left open, by Tait: the \( \Gamma^{(m)} \) principles are all inconsistent for \( m \geq 3 \).

So, even with this syntactical constraint these classes of reflection principles are either inconsistent or are still intra-constructible.

Koellner finishes his Section 4 with a heuristic argument that any form of reflection principle which is consistent relative to large cardinals is consistent relative to \( \kappa(\omega) \). If \( \kappa \) is \( \omega \)-Erdös, then \( (V_\kappa, \in) \) has an infinite sequence of indiscernibles \( I \subseteq \kappa \).

Take the Skolem hull \( H \) of \( I \) in \( V_\kappa \). Then any order preserving map \( j_0 : I \to I \) induces a non-identity first order elementary map \( j : M \to M \), where the ZFC-model \( M \) is the transitive collapse of (the countable) \( H \). We then have a situation similar to the (inconsistent with AC) assertion that there is a non-trivial elementary \( j : V \to V \). Koellner argues that from the point of view of consistency proof “it would appear that whatever reflection is provable from \( j : V \to V \) should also be provable from \( j : M \to M \). Since reflection would appear to be an entirely internal matter, this is a reason for thinking that any conceivable reflection principle must have consistency strength below that of \( \kappa(\omega) \).” (My emphasis) Well, is reflection an entirely internal matter? The view I shall be putting forward here, is that it is not. It is, or can be widened to be, a Gesamtauffassung that incorporates the whole, consisting of both the Cantorian sets and absolute in\( \to \)finites. If so, then it is not internal, and we have a hope for finding extraconstructible principles.

4. The Ackermannian realm and reflection

Another set theory, due to Ackermann was introduced \([1]\) and studied in the 1950’s and 60’s. Ackermann’s set theory \( A \) provided for a universe with extensionally determined entities (classes) and a predicate \( \bar{V} \) for set-hood: “\( x \in V \)”. Besides axioms for extensionality, a class construction scheme, and set completeness (“all classes that are subclasses of sets are sets”), it contained the following crucial principle:

- **(Ackermann’s Main Principle)** If \( X \subseteq V \) is definable using only set parameters, and not using the predicate \( \bar{V} \), then \( X \in V \). Thus if \( \theta \) does not contain \( V \):

\[
x \in V \land \forall t (\theta(x, t) \rightarrow t \in V) \quad \rightarrow \quad \exists z \in V \forall t (t \in z \leftrightarrow \theta(x, t))
\]

Ackermann interpreted Cantor’s “By a set we understand any collection of definite distinct objects … into a whole” as saying “we must require from already defined sets that they are determined and well-differentiated, thus the [foregoing] conditions for a totality [to be a set] only turn on that it must be sufficiently
sharply delimited what belongs to a totality and what does not belong to it. However now the concept of set is thoroughly open.”

(Ackermann [1] p.337)

Indeed Reinhardt, whilst working from the premise that Ackermann considered the concept of set itself as not sharply delimited ([18], p190-1), surmises that the intuition behind Ackermann’s Main Principle is that a sharply defined collection of sets is a set, and that, given the set \(x\), the property ‘\(t\) is a set such that \(\theta(x,t)\)’ is independent of the (extension of) the concept of set, but gives a sufficient condition for a collection to be sharply delimited. We therefore see that on the other hand a collection is not sufficiently well-differentiated if it is defined through its relationship to the concept of set.

The Ackermann quotation continues (in paraphrase) that in the Cantorian definition it is intended that a collection should be investigated only on a case by case basis as to whether it represents a set, and it is not meant that it is determined all at once for all classes whether they are sets or not.

Levy, Vaught [12] added Foundation to \(A\) calling this \(A^*\). Then \(A^*\) is consistent relative to \(A\), and proves the existence of the classes: \(\{V\}\), \(\{\{V\}\}\), \(\mathcal{P}(V)\), \(\mathcal{P}(\mathcal{P}(V))\)...

Thus the classes over \(V\) in \(A^*\) advance for infinitely many types beyond. The picture is thus quite different from a first-order, or even second-order ZF. However, Levy considered models of \(A^*\) of the form \((V_\alpha,\in, V_\beta)\). The paper [11] showed that \(A^*\) is \(\mathcal{L}_{\in_2}\)-conservative over ZF: \(A^* \vdash \sigma^V \implies ZF \vdash \sigma\).

Reinhardt in [18] proved the converse implication of this last result also; hence putting these together the set-theoretical content of \(A^*\) had always just been that of ZF. It is thus to be noted that two rather different conceptualisations - the one leading to the ZF formalisation, the other that of Ackermann’s - have the same content as far as the strictly considered set part is concerned. Reinhardt considered in [19] (and [20]) ideas that involved having an ‘imaginary realm’ beyond the Cantorian universe \(V\) which he wrote as \(V_\Omega\). He imagines having classes, say \(P\), which are then ‘projected’ into the imaginary realm as \(jP\). The difference between classes and sets is that the projection of the latter are themselves, whilst that of one of the former contains more imaginary sets and ordinals. \(V_\Omega\) itself is an imaginary set in this projected universe. Going yet further, he imagines a typed hierarchy of ‘\(\Omega\)-classes’ up to some \(\lambda > \Omega\), and collecting these together as \(V_\lambda\) he will project \(V_\lambda\) into some virtual realm \(V_{\lambda'}\). He formulates an extendability principle \(E_0(\Omega, \lambda; \Omega, \lambda')\):

(i) \(\Omega < j\Omega = \Omega' < \lambda'\).
(ii) \(\forall x \in V_\Omega jx = x\);  
(iii) \(j : (V_\lambda, \in) \rightarrow_{\Sigma_0} (V_{\lambda'}, \in)\).

Of course this has come down to us distilled in set-sized form as heading towards the definition of \(\alpha\)-extendible cardinals. In all of these theories, there is formed the idea of some ‘realm’, ‘universe’ etc. beyond \(V\).

We mention these Reinhardtian views of upwards projection of \((V, \in)\) by way of a contrast to the Global Reflection Principle to come.

\[^3\]... wir von den schon definierten Mengen verlangen müssen, dass sie bestimmt und wohlunterschieden sind, so kann es sich bei der obigen Bedingung für eine Gesamtheit nur darum handeln, dass genügend scharf abgegrenzt sein muss, was zu der Gesamtheit gehört und was nicht zu ihr gehört. Nun ist aber der Mengenbegriff durchaus offen."
5. Global Reflection

If we are contemplating an ab initio conception of the universe of sets, then we may proceed as follows.

By “conception of the universe of sets” we mean here something like the notion of “concept of set structure” in one of Martin’s versions of concept of set [16]. He writes that for him the modern, iterative concept has four important components:

1. The concept of extensionality
2. Concept of ‘set of $x$’s’
3. Concept of transfinite iteration
4. Concept of absolute infinity.

He is thinking of the concept of sets as the concept of ‘structuralist’s structure’ and thus does not have to add anything as to what kind of things sets are. We adopt this view here. (Martin remains silent as to which flavour of structuralism’s structure might be at play here.) A ‘set structure’ is then what is obtained by iterating the concept ‘set of $x$’s’ absolutely infinitely many times.

Still again at some pre-formal stage, he then takes some Informal Axioms encapsulating set theoretical principles (Extensionality, Comprehension) and rehearses the categoricity arguments going back to Zermelo, that for any two $\mathfrak{V}_1 = (V_1, \in_1)$, $\mathfrak{V}_2 = (V_2, \in_2)$ obtained by iterating the models’ $V_\alpha$ function throughout all the absolute infinity of ordinals, we have an isomorphism $\pi : (V_1, \in_1) \rightarrow (V_2, \in_2)$. In short, whatever view we take of what exact set formation process takes place when we take the “set of $x$’s”, we end up with isomorphic universes. We, as set theorists, thus shall pay no more attention to the mysteries of what exactly “the set of $x$’s” is or what precisely $\in$ is, than we do every day, and shall simply refer to the set theoretic universe as $(V, \in)$. But we do further remark for later that $\pi \upharpoonright \text{On}^{\mathfrak{V}_1} : \text{On}^{\mathfrak{V}_1} \cong \text{On}^{\mathfrak{V}_2}$ where $\text{On}^{\mathfrak{V}_i}$ is the absolute infinity of von Neumann ordinals in the structure $\mathfrak{V}_i$.

We then proceed in his paper as follows. We consider the universe of sets, $V$, (as above, unique up to this informal isomorphism argument) as the universe of the domain of purely mathematical discourse: whatever mathematical objects the mathematician needs, there are (isomorphic copies of) such in $V$. Indeed we regard sets themselves as mathematical objects. As we know, of necessity there are entities outside of $V$, where the modality of ‘necessity’ here is ‘logical necessity’: logic requires that the Russell class, or the class of ordinals, or indeed $V$ itself is not a set. We swallow the Cantorian pill that there are two kinds of entities: the mathematical-discourse sets, and the absolute infinities.

However we depart from von Neumann, who seemingly treated both kinds of entities in an equally ‘mathematical’ spirit (see [22], [23]) when developing his functions-as-classes theory: his classes were subject to mathematical laws. We draw a firmer conceptual line, and do not treat the absolute infinities in such a mathematico-functional manner.

In a paper with Leon Horsten ([25])¹ we have recently discussed the possible interpretations of classes prior to a development of a formal theory of them. We rule out a theory of classes as plurals: a plural, which in any case is a linguistic construct, is not supposed to add anything more ontologically to the objects we

---

¹Readers of which will note how much of the discussion here is indebted to it. I want to express my thanks to him for his initial suggestion for that paper and our many fruitful, and at times very entertaining, discussions on it.
have, namely sets. However we do have more, $V$ is not just “some sets so that $x = x$”. We accept that classes are entities that enter into structural arguments, which do not have to have any *prima facie* description as a plurality. However a mereological description of absolute infinities, as being parts of $V$, the absolute infinity of all sets, allows us to give sufficient substance to these entities without tying them to any language, or syntax. We may take over a theory of mereology, such as Lewis [13],[14] and apply this to $V$ together with its parts. (We have to make some adjustments: Lewis is dubious about $\emptyset$ and the $x \rightarrow \{x\}$ map; but we shall ignore these and treat our theory of parts as one which identifies sets $x$ also as ‘small’ parts.) Thus we take sets and the set elementhood relation as a given: we are not trying to alter our conception of sets. Lewis thinks that the parthood relation goes someway to help us understand the set-elementhood relation, but we are not committing ourselves to what exactly this latter relation is.

We may thus think of second order quantifiers, if later we come to formalise our notions, as ranging over the parts of $V$. It should be not associating the parts of $V$ with any particular linguistic structure: whereas pluralities could be interpreted in some minimal NBG model sitting above $(V, \in)$, a mereological view can sit happily with a Kelley-Morse formal theory of classes, but on its own is not restricting the absolute infinities that there are. How could it?

One should also note that there is no hierarchy of “super-parts” of collections or “collections of parts” or any such that threatens to build a ramified hierarchy of classes beyond On: the power set operation that collects together all the subsets of a set into a set is a mathematical operation applied to sets. Our acceptance of a power set operation does not require us to countenance a “power-absolute-infinity” operation. To insist that we must consider such an operation, if we posit it for sets, is similar to insisting that if our (physical, space-time) universe is finite then “there must be something beyond it”.

We denote by $\mathcal{C}$ the collection of parts of $V$. We identify parts of $V$ that are parts of sets, as themselves sets. The other parts of $V$ are the absolute infinities. Then $(V, \in, \mathcal{C})$ is the realm of ‘Cantorian discourse’. Admittedly $\mathcal{C}$ inherits the ineffability of the notion of absolute infinity. Initially then $\mathcal{C}$ would have been populated with examples of absolute infinities that we are familiar with and had been defined by the early researchers. But we do not insist on restricting to this definability. (We cash in the remark we made earlier that two possible notions of “set of $x$’s” led to isomorphic universes $\mathcal{U}_1 = (V_1, \in_1)$, $\mathcal{U}_2 = (V_2, \in_2)$ with an isomorphism $\pi$ between them, which in particular restricted to an isomorphic map between the absolute infinity of $\mathcal{U}_1$’s ordinals with those of $\mathcal{U}_2$. The same argument shows that ‘parts of $\mathcal{U}_1$’ carry over to parts of $\mathcal{U}_2$. So clearly we may extend the isomorphism to $\pi : (V_1, \in_1, \mathcal{C}_1) \rightarrow (V_2, \in_2, \mathcal{C}_2)$. In other words two differing notions of “set of” cannot lead to essentially differing models when their parts are also included. Moreover we view the content of our set, and class, theoretical ideas to be captured by this isomorphism type.)

We want to invert the Reinhardtian approach and stand it on its head: instead of projecting into some ‘virtual realm’, we reflect the structure $(V, \in, \mathcal{C})$ to some set-sized part of itself. The approach taken is that we regard $(V, \in, \mathcal{C})$ as absolutely indistinguishable from one of its initial segments.
**Definition 1. (GRP)** The Global Reflection Property holds if the universe $(V, \in, \mathcal{C})$ admits of a set-sized $\mathcal{L}_\mathcal{E}^+$-elementary substructure, $(V', \in, \mathcal{C}') \prec (V, \in, \mathcal{C})$ with $V' = V_\kappa$ for some $\kappa$ and which is isomorphic to a structure of exactly the same kind, namely $(V_\kappa, \in, V_{\kappa+1})$.

We thus intend the principle as saying that there is a set-sized version of $(V, \in, \mathcal{C})$ (that is $V$ together with its classes) which is of exactly the same kind, namely is some $(V_\kappa, \in, V_{\kappa+1})$ (so $V_\kappa$ together with its classes) and with some (restricted) elementarity reflected between them. Here $\mathcal{L}_\mathcal{E}^+$ denotes the usual first order language $\mathcal{L}_\mathcal{E}$, but augmented with second order free variable symbols $\bar{A}, \bar{B}, \ldots$. The interpretation of the second order variables to range over the collections $\mathcal{C}, \mathcal{C}'$ of parts of $V$. The second order variables are thus not quantified over. (We shall write, for example, ‘$\Sigma^*_n$’ for formulae at that level of complexity in $\mathcal{L}_\mathcal{E}^+$, to indicate $n$-alternating set quantifiers.)

It is also possible to see this indistinguishability as the endpoint of a spectrum of more limited reflection properties which we outline first. Just for convenience we phrase these as embedding properties more familiar to set theorists.

**Definition 2. (Limited Global Reflection)** There is $\kappa \in \operatorname{On}$, there is $j \neq \text{id}$, with $\operatorname{crit}(j) = \kappa$, and a non-empty $D \subseteq V_{\kappa+1}$ so that:

$$j : (V_\kappa, \in, D) \rightarrow \Sigma^*_2 (V, \in, \mathcal{C}).$$

Here we stipulate that $\rightarrow_{\Sigma^*_2}$ denotes an embedding that preserves truth of $\mathcal{L}_\mathcal{E}^+$ formulae. As $\operatorname{crit}(j) = \kappa$ we have: (i) $j|V_\kappa = \text{id}|V_\kappa$; (ii) if $\kappa \in D$ then $j(\kappa) = \operatorname{On} \in \mathcal{C}$. We thus have:

$$\varphi(x, X)^{(V_\kappa, \in, D)} \iff \varphi(x, j(X))^{(V, \in, \mathcal{C})}$$

The strength of such a principle depends entirely on the nature of $D$, which is the non-trivial part of the domain of $j$. If $D$ is extremely thin, then the property is saying hardly anything at all. If we had allowed $D = \emptyset$ then all we should have had is that $(V_\kappa, \in) \prec \Sigma^*_2 (V, \in)$.

- If $D \subset P(\kappa)^L$ then in general we shall not have a reflection property that is extra-constructible. Indeed, several ‘classical’ reflection properties can be expressed in this way.
- However if $D \supseteq P(\kappa)^L$ then we can define in the usual way an $L$-ultrafilter on $\kappa$:

$$X \in U \iff (X \in L \land \kappa \in j(X)) \quad (1)$$

By standard arguments this is an amenable normal ultrafilter, and we may define a wellfounded ultrapower $\operatorname{Ult}((L, \in), U)$ which is isomorphic to $L$ itself. In other words, we have some non-identity embedding $j_0 : L \rightarrow L$, i.e. $\emptyset$ exists.

- As $D$ is reckoned to be larger and larger, then the strength of the principle increases: if some other definable inner model $M$ has $D \supseteq P(\kappa)^M$ then again we shall be able to define an ultrapower of $M$: $\operatorname{Ult}((M, \in), U)$ if $U$ is defined in the same way. Such a model may then also be seen to be non-rigid.

The logical limit, and principle of main interest here, is when $D$ becomes maximal at the end of this spectrum, i.e. to become $V_{\kappa+1}$. Unlike the pitfall that was waiting for Reinhardt, this principle when extended to the limit can be shown consistent relative to large cardinals.
We then have the following GRP \textit{paraphrase}. \footnote{I am indebted for helpful discussions with Sam Roberts. In Def.1 above we define \textit{ran} \((j)\) and define from it \(j\) as the inverse of the transitive collapse. This is perhaps merely presentational: either way one still has to posit something whether the whole isomorphism \(j\) or what amounts to the same thing, its range, \(\text{ran}(j)\). Roberts has proposed principles extending GRP that posit (under certain conditions) in effect the range of a \(j\) (rather than \(j\) itself) as the result of a \textquote{Reflection process} in which one has an auxiliary class satisfaction principle.}

There is \(\kappa \in \text{On},\) there is \(j \neq \text{id},\) with \(\text{crit}(j) = \kappa,\) so that:

\(j : (V_\kappa, \epsilon, V_{\kappa+1}) \rightarrow \Sigma_\kappa (V, \epsilon, \mathcal{C}).\)

Some points are then clear: \(\kappa\) is strongly inaccessible; as there is a wellordering \(W\) of \(V_\kappa\) in \(V_{\kappa+1},\) then \(j(W)\) is a wellordering of \(V.\) Thus we must have global choice holding in \(V.\)

\(\bullet\) GRP is equivalent to the principle obtained by weakening \(\Sigma_\omega^0\) by \(\Sigma_0^0\) (but not by the usual self-strengthening argument of Gaifman, because that requires the range of the map \(j\) to be cofinal - which does not apply here).

As soon as we have \(D = V_{\kappa+1},\) we have that the ultrafilter \(U\) defined at (1) by removing the conjunct \(X \in L\) is a normal measure in \(V.\) Hence \(\kappa\) is measurable, and by the supposed elementarity, and by simple reflection arguments, we immediately have a proper class of measurable cardinals. But we easily have more.

\textbf{Theorem 1.} GRP implies there is a proper class of Shelah cardinals.

\textbf{Proof:} Recall that \(\mu\) is Shelah if \(\forall f \in {}^\mu \mathcal{N}, j (j : V \rightarrow N \land V_{j(f)(\mu)} \subseteq N).\) We show that \(\kappa\) is Shelah in the statement of GRP and this easily implies by elementarity that there is a proper class of such.

Let \(f \in {}^\kappa \mathcal{V}_{\kappa+1},\) be arbitrary. Then \(j(f) : \text{On} \rightarrow \text{On} \text{ and } j(f)^* \subseteq \kappa.\)

Take \(\lambda > \kappa\) a sufficiently large inaccessible, so that \(j(f)(\kappa) < \lambda,\) and consider the \(\lambda\)-strong extender derived from \(j:\)

For \(a \in [\lambda]^{<\omega}: E_a = \{z \in \mathcal{P}([\kappa]^{|a|}) : a \in j(z)\}; \quad \mathcal{E} = \langle E_a : a \in [\lambda]^{<\omega} \rangle.\)

This has the following properties:

\(1)\) \(\mathcal{E}\) is a \((\kappa, \lambda)\)-extender such that \(\text{Ult}((V, \epsilon), \mathcal{E})\) is wellfounded; with \(k_\mathcal{E} : V \rightarrow \text{Ult}((V, \epsilon), \mathcal{E}),\) and if \(l : \text{Ult}((V, \epsilon), \mathcal{E}) \cong N,\) is the unique transitive collapse map, then setting \(j_\mathcal{E} = l \circ k_\mathcal{E}, \quad j_\mathcal{E} : V \rightarrow N \text{ and } j(f)(\kappa) = j_\mathcal{E}(f)(\kappa) < \lambda,\) and \(V_\lambda = (V_\lambda)^N.\)

As \(j_\mathcal{E}(f)(\kappa) = k_\mathcal{E}(f)(\kappa) < \lambda,\) then \(V_{j_\mathcal{E}(f)(\kappa)} \subseteq N\) follows; hence we have the Shelah property for this \(f.\)

Clearly stronger properties can be argued for, but in any case we have \(\text{UW}:\)

\textbf{Corollary 1.} GRP implies there is a proper class of measurable Woodin cardinals.

\textbf{6. Is GRP a reflection principle?}

Instead of viewing GRP as a limiting principle as the class domain grows larger until it contains all of \(P(\kappa),\) one could view it outright as asserting in a strong form that \((V, \epsilon, \mathcal{C}),\) that is \(V\) \textit{together with} all its parts, is indistinguishable from one of its
initial segments and its parts: \((V_\kappa, \in, V_{\kappa+1})\). We view \(V\) together with its parts, and its initial segments and their parts as being so rich that we may form a substructure \((V', \in, C')\) with \(V' = V_\kappa\) for some \(\kappa\) such that \((V, \in, C)\) and \((V_\kappa, \in, V_{\kappa+1})\) can stand in this relationship. The latter is a simulacrum of the former. The inverse of the isomorphism \(j\) mediates that indistinguishability through being a truth preserving elementary embedding.

If this viewpoint is viable, then we are taking the whole of \((V, \in, C)\) and reflecting this to a \((V_\kappa, \in, V_{\kappa+1})\). It is not a syntactic, formula-by-formula reflection, whether first or second order, or something expressed in some logic. In these senses it is not a reflection principle such as Gödel may have had in mind. It is not the viewpoint that says “nothing we can say in whatever logic/language pins down \((V, \in, C)\), or is only true in \((V, \in, C)\)” (such a viewpoint would be too weak for our purposes). To assert GRP is to assert that there is a substructure of the given kind, or that there exists an isomorphism \(j\) doing the work of linking these collections of parts together, those of \(V_\kappa\) and those of \(V\). One cannot argue for GRP using the iterative conception of set alone, that is for a concept of ‘set-structure’ (with the components (1)-(4) listed at the beginning of the last section), but must be coupled with the reflective idea involving \(C\).

Of course the assertion of the existence of the substructure \((V_\kappa, \in, C')\) (or equivalently the interceding map \(j\)), is a second order existential assertion, and by elementary arguments, it cannot be a definable class of \((V, \in, C)\). So when we come to formalise our principle GRP this will require the admission that impredicative objects such as \(C'\) or \(j\) should be part of the discussion. There is no sense that the \(j\) of GRP has ‘come from somewhere’ or is ‘canonical’ in any way (it certainly cannot be definable). Friedman and Honzik ([6], Sect. 2) raise this non-canonicity as somehow a defect of GRP. “However, in our opinion, such strong forms of reflection seem to be too “uncanonical” to count as true formalization of (Reflection)”.

However their paper is concerned with something quite different from reflection of an ‘actualised’ universe \((V, \in)\); it is not a potentialist version either. It seeks to gain some insight into possible new axioms or hypotheses (such as the “Inner Model Hypothesis” that may then be offered for consideration as true of \(V\), by looking at countable transitive models, pretending that \(V\) is one of them, and thereby possibly “sharp-generating” this c.t.m. by iteration of an external premouse. Hence our viewpoint and this procedure are quite different (and at odds with each other). Perhaps the externally provided iteration maps \(\tilde{\pi}\) when restricted to the c.t.m. \(V\) are deemed more ‘canonical’? (Although, if so, one may counter this canonicity, by remarking that the externally posited ‘premouse’, and so its maps, are not unique either.) However this is a rather sophisticated approach that applies to countable transitive models ‘\(V\)’, and does not really touch (in this author’s view) a Cantorian view of \(V\) as an inhaustable totality of sets together with a pre-formal conception of reflection.

Indeed Peter Koellner has suggested that perhaps a “resemblance” property is a better name. Whatever one thinks of the nomenclature, GRP is different from other properties that are sometimes termed resemblance. We consider some such forms now for a differential comparison; these are forms which assert that there is some reflection or accumulation point \(\kappa\) so that anything that occurs above \(\kappa\) in some sense has an occurrence below. One such is the Vopenka principle:
Definition 3. (Vopenka’s Principle) If \(<M_\alpha | \alpha \in \text{On}>\) is a proper class of first order structures each in \(V\) of the same signature, then there is \(\alpha < \beta\) and an elementary embedding \(j : M_\alpha \rightarrow M_\beta\).

Notice that this is a richness principle of \(V\) that claims of any such class of structures that we have such a triple \(\alpha, \beta, j\). We can tie it closer to the structure of \(V\) by taking the \(M_\alpha\) of the form \((V, f(\alpha), \in, \{\alpha\}, R_\alpha)\) where \(\alpha < f(\alpha)\) for some increasing function \(f : \text{On} \rightarrow \text{On}\); and where \(R_\alpha \subseteq V_{f(\alpha)}\). The presence of \(\{\alpha\}\) makes all the difference as then if we have \(\alpha, \beta, j\) as in Vopenka’s Principle, then we must have \(j(\alpha) = \beta\), and thus \(j\) is not simple identity. Hence VP is a strong property which implies that the class of extendible cardinals is Mahlo in \(\text{On}\) (see [17]).

A more extensive study of VP like principles where the target structures have certain elementarity properties in \(V\) is given in Joan Bagaria’s paper [2]:

Definition 4. (i) \(C^{(n)} = \{ \alpha | (V_\alpha, \in) \prec_{\Sigma_n} (V, \in) \}\);

(ii) \(\kappa\) is a \(C^{(n)}\)-measurable cardinal if it is the critical point of an elementary embedding \(j : V \rightarrow M\), with \(M\) transitive, and \(j(\kappa) \in C^{(n)}\);

(iii) \(\kappa\) is a \(C^{(n)}\)-extendible cardinal if for all \(\lambda\) there exists \(\mu, j\) with \(\text{crit}(j) = \kappa\); \(j : V_\lambda \rightarrow_{\Sigma_n} V_\mu\) with \(j(\kappa) \in C^{(n)}\).

As [2] analyzes, if \(\kappa\) is measurable then it is \(C^{(n)}\)-measurable, and the prefix in this sense adds nothing; whilst, \(C^{(n)}\)-extendibility is a genuine strengthening of extendibility.

One final definition before we can state Bagaria’s categorization. This is now a form of resemblance where \(\kappa\) is some kind of ‘reflection point’ or ‘cut point’ in the universe \(V\):

Definition 5. \(\text{VP}(\kappa, \Sigma_n)\) holds iff for every proper class \(C\) of structures of the same type \(\tau\) such that both \(\tau\) and the parameters of some \(\Sigma_n\)-definition of \(C\), if any, belong to \(H_\kappa\), then \(C\) reflects below \(\kappa\), i.e.,

\[\forall B \in C \exists A \in C \cap H_\kappa (A \text{ is elementarily embeddable into } B)\]

Notice this is not so dissimilar structurally to the GRP: to assert the existence of an elementary embeddable \(A\) as above, of course is also to assert the existence of an elementary substructure of \(B\), namely the range of that elementary embedding, which is isomorphic to a structure of the same kind: the original \(A\) in \(C\). GRP also asserts the existence of a substructure isomorphic to a small (equals set-sized) one of the same kind.

Theorem 2. (Bagaria [2], 4.15) The following are equivalent:

(i) \(\text{VP}\);

(ii) For all \(n\) there is a proper class of \(\kappa\) so that \(\text{VP}(\kappa, \Sigma_n)\);

(iii) \(\forall \exists \kappa (\kappa\text{ is }C^{(n)}\text{-extendible})\).

We thus see that VP can be given an equivalence in terms of a proper class of reflection or cut points for any kind of definable class whatsoever.

This in fact brings out parallels with a much earlier paper of Magidor’s on supercompacts:

Theorem 3. (Magidor [17]) Let \(\kappa\) be the first supercompact cardinal; then \(\kappa\) is the least cardinal so that \(\text{VP}(\kappa, \Sigma_2)\).
We have mentioned the details of these definitions to see that the kind of reflection they represent is some form of internal reflection of the ramified layers $V_\alpha$ rather than the idea of reflection of the whole of universe $(V, \in)$ which cannot be pinned down in ways mentioned by Ackermann and Gödel. So there is a qualitative difference between GRP and such principles.

However it is easily noted that:

**Theorem 4.** $\text{Con}(\text{ZFC} \vdash \exists \kappa (\kappa \text{ is } 1\text{-extendible})) \rightarrow \text{Con}(\text{NBG} + \text{GRP}).$

But the arrow is not reversible. GRP thus falls just sort of implying the existence of those embeddings $j$ that are discontinuous at the successor of the critical point: $j^\kappa$ is bounded in $j(\kappa^+)$. Hence GRP is consistent, or can be made consistent, with Global Square, and $\square_\lambda$ everywhere, by a class forcing. (For these methods see [3] It is thus a reflection principle that marks off this threshold.

## 7. Strengthening GRP?

Whilst the last theorem indicates what the strength of the basic GRP is, the motivation for top-down reflection rather than upwards projection came originally from a weakening of the notion of subcompactness:

**Definition 6.** $\kappa$ is subcompact if for any $A \subseteq H_{\kappa^+}$, there are $j, \mu < \kappa$, with $\text{crit}(j) = \mu, j(\mu) = \kappa$ and a $B \subseteq H_{\mu^+}$ with:

$$j : (H_{\mu^+}, B) \rightarrow^c (H_{\kappa^+}, A).$$

To strengthen GRP we may ask for many $j$’s and $\kappa$’s. Or else, more interestingly, we may increase the elementarity of $j$ to be (partially or fully) second order reflecting, that is, to preserve for example $\Sigma^1_\epsilon$-formulae with now quantification over the second-order variables of $L^+_{\epsilon}$ - such an extended language we shall call $L^2_{\epsilon}$. Then if $j : (V, \epsilon, V_{\kappa+1}) \rightarrow^c (V, \epsilon, \mathcal{C})$ we shall conclude that all instances of impredicative comprehension - which are true in $(V, \epsilon, V_{\kappa+1})$, being a Kelley-Morse (KM) model, will also hold in $(V, \epsilon, \mathcal{C})$.

If $j$ is $\Sigma^1_\epsilon$-elementary more can be said about the range of $j$: for example $\Sigma^1_\epsilon$-elementarity shows that the class of Shelah cardinals is stationary in $(V, \epsilon, \mathcal{C})$. We may go further and formalise second order satisfaction, as follows.

We may for any $n \in \omega$ define a $\Sigma^1_n$ formula $\text{Sat}_n(v_0, \ldots, v_k, Y_0)$ so that, provably in NBG +Global Choice (the latter holds if GRP does, and is needed to define Skolem functions):

$$\forall h \in ^n_v \forall X_1, \ldots \forall X_m$$

$$\text{Sat}_n(\langle \varphi \rangle, k, m, \langle h_0, \ldots, h_{k-1}, X_1, \ldots, X_m \rangle) \leftrightarrow \varphi(\overline{h}, X_1, \ldots, X_m),$$

for any $\Sigma^1_n$ formula $\varphi(v_1, \ldots, v_k, Y_1, \ldots, Y_m)$ with the $v_i$ to be interpreted as sets, and the $Y_i$ as classes. Let Sat be the amalgamation of these Sat$_n$ predicates.

**Definition 7.** $(\text{GRP}^+)$. There is $\kappa \in \text{On}$, there is $j \neq \text{id}$, $\text{crit}(j) = \kappa$,

$$j : (V, \epsilon, V_{\kappa+1}, \text{Sat}^{V_\kappa}) \rightarrow^c (V, \epsilon, \mathcal{C}, \text{Sat}).$$

We thus require $j$ to be $\Sigma^1_\epsilon$-elementary in the full second order language $L^2_{\epsilon, \overline{S}}$ with a predicate $\overline{S}$ for Sat. It is easy to argue that subcompactness of some $\lambda$ then yields a model of $\text{GRP}^+$. Just to demonstrate the additional strength of this principle over GRP we show the following.
Proposition 1. Assume GRP\(^+\). Then there is a commuting system
\[\langle \mu_\alpha, j_\alpha \rangle_{\alpha \leq \beta \in \text{On}}\]
of embeddings \(j_\alpha : (V_{\mu_\alpha}, \in, V_{\mu_\alpha+1}) \to \mathcal{V}_1\) \(V_{\mu_\beta}, \in, V_{\mu_\beta+1}\), with each \(j_\alpha, \alpha < \beta\), witnessing the simple GRP at \(\mu_\beta\). Thus each \(j_\alpha|\mu_\alpha = \text{id}|\mu_\alpha\) and \(j_\alpha(\mu_\alpha) = \mu_\beta\).
Moreover for \(\alpha \in \text{On}\), there are maps \(j_\alpha : (V_{\mu_\alpha}, \in, V_{\mu_\alpha+1}) \to \mathcal{V}_1(\text{V}, \in, \mathcal{C})\) also witnessing GRP in the universe.

This is obtained in a way very similar to the following from a subcompact.

Proposition 2. Let \(\kappa\) be subcompact. Then there is a commuting system
\[\langle \mu_\alpha, j_\alpha \rangle_{\alpha \leq \beta \leq \kappa}\]
of embeddings \(j_\alpha : (V_{\mu_\alpha+1}, \in, V_{\mu_\alpha+1}) \to \mathcal{V}_1(\text{V}_{\mu_\beta}, \in, V_{\mu_\beta+1})\) with \(\mu_\alpha < \mu_\beta\) for \(\alpha < \beta; \mu_\kappa = \kappa\); each \(j_\alpha|\mu_\alpha = \text{id}|\mu_\alpha\), and \(j_\alpha(\mu_\alpha) = \mu_\beta\), and thus with each \(j_\alpha\), \(\alpha < \beta\), witnessing the \(1\)-extendibility of \(\mu_\alpha\).

Proof: For \(\lambda \in \text{Card}\) let \(\text{Sat}_\lambda\) be the satisfaction relation for \((H_\lambda, \in)\). Then we view \(\text{Sat}_\lambda\) as a subset of \(H_\lambda\). Let \(\kappa\) be subcompact as above, and let \(A = \text{Sat} = \text{Sat}_\kappa^+\). Then \(\text{Sat} \subseteq H_\kappa^+\), and applying the definition of subcompactness there are \(\mu < \kappa\) and \(j\), and \(\text{Sat}\) with
\[j : (H_\mu^+ \text{Sat}) \to (H_\kappa^+, \text{Sat}).\]

(1) \(\text{Sat} = \text{Sat}_{\mu^+}\).

Pf: Suppose \(H_\mu^+ \models \varphi(x) \leftrightarrow \neg \text{Sat}(\langle \varphi^\gamma, x \rangle)\). Apply \(j\) to get a contradiction. \(\Box(1)\)

Definition 8. \((H_\kappa^+, \text{Sat}) \models \langle \text{ran}(k) \prec_e (V, \in) \rangle^e_e\leftrightarrow ^e_e\)
\[
\forall x \in \text{ran}(k) \forall \varphi^\gamma \equiv \exists z \psi(z, v_1^\gamma) \exists z \text{Sat}(\langle \varphi^\gamma, (z, k(x)) \rangle) \implies \exists z \in \text{ran}(k) \text{Sat}(\langle \varphi^\gamma, (z, k(x)) \rangle).
\]

We thus use the standard Tarski-Vaught criterion on elementary substructure to formalise the notion of being an elementary submodel of \(V\). Note that \(j \subseteq H_\mu^+ \times H_\kappa^+\) and \(|j| = |H_\mu^+| < \kappa\) and so \(j \in H_\kappa^+\). Clearly, by the elementarity of \(j\):

(2) \((H_\kappa^+, \text{Sat}) \models \langle \text{ran}(j) \prec_e (V, \in) \rangle^e_e\)

\((H_\kappa^+, \text{Sat}) \models \langle \text{There are } k, \kappa_0, \text{ with } k : (H_{\kappa^+_0}, \in) \to (V, \in) \text{ and } \text{ran}(k) \prec_e (V, \in), \text{ and thus } k \text{ is an elementary map.} \rangle^e_e\)

But by invoking \(j\) we have:

(3) \((H_\mu^+, \text{Sat}_{\mu^+}) \models \langle \text{There are } k, \kappa_0, \text{ with } k : (H_{\kappa^+_0}, \in) \to (V, \in) \text{ and } \text{ran}(k) \prec_e (V, \in), \text{ and thus } k \text{ is an elementary map.} \rangle^e_e\)

This gives us two links in a chain of models we are looking for in the Proposition. Suppose there are no chains of length \(\kappa\) of the kind sought. Let \(C = \langle \mu_\alpha, j_\alpha \rangle_{\alpha \leq \beta \leq \tau}\) be a maximal such commuting chain with the properties: (i) there is a final model \((V_{\mu_{\tau+1}}, \in)\), and (ii) with a final map \(j_\tau : (V_{\mu_\tau+1}, \in) \to (V_{\kappa+1}, \in)\). Suppose \(\tau < \kappa\) (for otherwise we are done). For each \(\alpha < \tau\) we have that \((H_{\kappa^+}, \text{Sat}) \models \langle \text{ran}(j_\tau \circ \text{ran}(j_\alpha) \prec_e (V, \in) \rangle^e_e\)
By elementarity of \( j \) we have in fact that, \( C \) is similarly a 1-extendible chain, but now that there is some \( \mathcal{F}_e \) such that, \( (V_{\mu+1}, \in) \to (V_{\mu+1}, \in) \), i.e., with the chain also maximal such, but with target \( (V_{\mu+1}, \in) \). This however is a contradiction since \( (H_{\mu+1}, \text{Sat}) \) now sees that \( C \) can be extended one more link (namely \( \text{via } \mathcal{F}_e \)), and this reflects into \( H_{\mu+1} \), so \( C \) is not a maximal chain going up to \( (V_{\mu+1}, \in) \).

\( \square \)

8. Conclusion

The strengthenings of GRP considered in the last section go beyond the kind of position outlined earlier: the purely mathematical objects reside in \( V \), it is the parts of \( V \) that form the proper classes of \( C \). In GRP we reflect, in as much that there should be an initial segment \( V_{\nu} \) together with its parts, that is a simulacrum of \( (V, \in, C) \), that is, it is isomorphic to a substructure of \( (V, \in, C) \). We classify \( \Sigma^0 \) statements as mathematical; the second order expressions quantifying over parts are mereological: these are about the parts of \( V \). However if the substructure of \( (V, \in, C) \) is supposed to be \textit{merologically elementary}, that is, if the elementarity is such that it is required to reflect structural statements or further commitments about the parts of \( V \) beyond those expressible in \( \mathcal{L}_\in^1 \) (or in other words, that the embedding \( j \) is more than first order respecting) then this can be construed as a step beyond the pure Cantorian picture we have argued for.

References


School of Mathematics, University of Bristol, & Isaac Newton Institute for Mathematical Sciences, Cambridge, UK

E-mail address: p.welch@bristol.ac.uk

URL: http://www.maths.bristol.ac.uk/~mapdw