Set Theory in the Foundation of Math; Internal Classes and External Sets

Leonid A. Levin* (https://www.cs.bu.edu/fac/Lnd)

Abstract

Usual math sets have special types: countable, compact, open, occasionally Borel, rarely projective, etc. Each such set is described by a single Set Theory formula with parameters unrelated to other formulas. Exotic expressions involving sets related to formulas of unlimited quantifiers height appear mostly in esoteric or foundational studies.

Recognizing internal to math (formula-specified) and external (based on parameters in those formulas) aspects of math objects greatly simplifies foundations. I postulate external sets (not internally specified, constituting the domain of variables) to be hereditarily countable and independent of formula-defined classes, i.e. with finite algorithmic information about them.

This allows to eliminate all non-integer quantifiers in Set Theory sentences. All with seemingly no need to change almost anything in mathematical papers, only to reinterpret some formalities.

1 Introductory Remarks: The Problem

I always wondered why math foundations as taken by logicians are so distant from the actual math practice. For instance, the cardinality theory – the heart of the ZFC Set Theory – is almost never used beyond figuring out which sets are countable and which are not. I see the culprit in the blurred distinction between two types of collections different in nature.

One type is pure ***classes***, defined by math properties. The really useful math properties go to very limited heights. Yet, provisions for open-ended hierarchies are motivated by the perceived need to handle objects that are not defined by readily envisioned properties.

Indeed, mathematics needs to handle external objects, not defined by any math properties at all, for instance, random sequences. These ***externals*** make the other type: math handles them but generally does not specify internally.[1] I take externals as proper sets forming the domain of Set Theory variables. I postulate them to be hereditarily countable and independent of pure classes, i.e. have only finite algorithmic information about them. Then, **math objects** (only informally called sets) are **classes** of sets specified by formulas with external parameters. Collections of objects are treated as collections of those parameters. Excluded are "Logics" objects involving formulas with unlimited quantifier heights, such as e.g., the collection of all true statements of Arithmetic.

Cantor's axioms asserted that all Set Theory formulas define (quantifiable) sets. In effect, this allows formulas with quantifiers over formulas. This self-referential aspect turned out fatal.

Zermelo and Fraenkel reduced this aspect by (somewhat *ad-hoc*) restrictions on cardinality treatment by Cantor's Axioms: Replacement Axiom preserves cardinality bounds of existing sets;

*Boston University, CAS, Computer Science department, 665 Commonwealth Avenue, Boston, MA 02215.

[1] In reality, external data end up finite. Yet, infinities are neat. Even for finite objects their termination points are often ambiguous and awkward to handle. As 0 is a great simplifying approximation to negligible ε, so is ∞ for $\frac{1}{\varepsilon}$. And note that the set $\overline{\mathbb{R}}$ of infinitely long reals is compact, having more "finite-like" features in that than $\overline{\mathbb{Q}}$.

only a separate Power Set Axiom increases them. This cardinality focus has questionable relevance. Distinctions between uncountable cardinalities almost never looked at in math papers.

And there is something exotic for mathematics in the Power Set axiom. Usual math sets have special types: countable, compact, open, occasionally Borel, rarely projective, etc. Generic subsets from Power Set classes, with no other descriptions, find little use in mathematics.
(Admittedly, the sets provided by rarely used axiom of Choice may pose an issue here. But even they can be, with some stretch, looked at as internally specified – by adding to the language a notation for (named, treated as canonical, but not described) injection of reals into ordinals.)

And as all sound axiom systems have countable models, cardinalities feel like artifacts, designed to hide self-referential aspects. Many papers in Logic (e.g., [2, 3, 7]) aimed at isolating math segments where more ingenious proofs can replace the use of Power Set Axiom and its uncountable sets. But this breaks the unity of math: an unfortunate effect. And complicating proofs is unattractive, too.

Some Formalities. In effect, math objects would have ranks, each quantifier binding only classes below one rank. Classes F^k_p of rank $\le k$ are specified by a hereditarily countable parameter p via the universal Σ^1_k formula $F^k_p(q)$, defining membership $F^k_q \in F^k_p$ on their transitive closures.

It is convenient to include in p a v.Neumann ordinal $o(p)$ and condition membership on $o(q){<}o(p)$. Then Foundation axiom needs to be stated only for classes of sets (ordinals). Membership also extends by extensionality: $F^{k'}_{q'}{\in}F^k_p$ if $F^{k'}_{q'}{=}F^k_q$, i.e. if membership relation graphs of transitive closures of $F^{k'}_{q'}$, F^k_q have a formula-defined isomorphism. Typical math concepts have straightforward translations. E.g., "x is in open $P \subset \mathbb{R}$" can be a shorthand for "$F^k_p(x)$, where p specifies $P \stackrel{\text{df}}{=} F^k_p$ as the set of its rational intervals". (Or P can be Borel, or whatever type clear from the context.)

2 Purging the Tricksome Objects

Making explicit the internal to math (defined by formulas) and external (parameters based) aspects of math objects clarifies their nature, allows a better focus on issues brought by these different (albeit with many similarities) sources. "Pure" classes defined by parameter-free Set Theory formulas are specific but tricksome. Treated carelessly they easily bring paradoxes. They will not be quantified or put in unlimited definable collections: that would only extend the language of allowed formulas.

Set Theory variables then range over "proper" sets: the external data. No reason to expect any self-referential issues from them. Such issues come from math properties on variables ranging over objects defined by properties themselves. This need not be the case with variables covering only external objects, unrelated to math properties. Externals may be chaotic, but not really tricksome.

Those sets can be put in collections based on properties and relations with other sets. This forms general math objects: classes, specified by Set Theory formulas with some free variables taken for external parameters. They carry both tricks and chaos. But tricks are based only on single formulas.

It is redundant for formulas to use in their parameters information from formulas themselves. This would not increase their expressive power. Besides, how could external parameters acquire infinite uncomputable information about formula-defined classes ? Yet, seeing no need or no realistic mechanisms for such occurrences is not the same as ruling them out.

Here the algorithmic information theory comes to the rescue. (See sec. 3.1 for some background.) With sets clearly distinguished from pure classes it allows a radical insight. Parameters may have not only infinite size (like, say, the number π), but also infinite complexity. Yet, even with infinite complexity, external sequences can have only finite dependence on pure classes. This is formalized as them having finite (small, really) mutual algorithmic information.

The "physical" meaning of that stems from Independence Conservation Inequalities. They state that no algorithmic or random processes, or any their combinations, can increase the amount of algorithmic information their inputs have about any specific sequence, such as one defined by a formula. See [5] for a more detailed discussion.

This justifies a family of axioms that, for any pure class of integers, prohibits external sequences to have infinite information about that class. This powerful family, called Independence Postulate (**IP**), opens a way to reduce any set theory sentence to one with only integer quantifiers.

Note, some sequences are both generatable (thus legitimate externals) and definable by formulas. But by **IP** they all are computable, otherwise they would have infinite information about themselves (a sort of Church's Thesis). So, **IP** conflicts with Replacement axioms for sets.
(For classes, these axioms are merely definitions.) Thus I restrict Replacement axioms to just one: Classes whose membership graphs are ***computably enumerable (c.e.)*** **from a set** are sets.

Note, **IP** excludes $\alpha \in F \in \Delta^0_* \stackrel{\rm df}{=} \cup_n \Delta^0_n$ unless such $\alpha \in \mathbf{\Omega} \stackrel{\rm df}{=} \{0,1\}^{\mathbb{N}}$ reduce to a positive fraction of all sequences. This condition involves only integer quantifiers. But the reverse does not follow from **IP**. So, it needs support from more axioms. One is the Primal Chaos axiom ($\mathbf{P}\boldsymbol{\mathcal{X}}$): "Each set reduces to (its membership graph is enumerable from) even-indexed digits of some Kolmogorov–Martin-Lof (**K-ML**) random $\beta \in \mathbf{\Omega}$." (For classes, i.e. in ZFC, it is Gacs-Kucera theorem: see [6].)

Note that a random sequence respects the Independence Postulate if and only if it is ***generic***, meaning that it is outside of all arithmetic classes $X \in \Delta^0_*$ of measure zero.

Our axioms are consistent, having, as all consistent theories do, a countable model.
Each model is directed under reducibility. This means that with each pair of sequences it includes a sequence they both are computable from and also all sequences computable from them.

A countable model has a reduction basis: a chain of sequences, each computable from the next one, and all sequences in the model are computable from them. By the axioms discussed, each of them reduces to a generic sequence (and can be taken Turing-equivalent to them).

We can represent such basis $\gamma^{(k)}$ as a set of trims of a combined sequence γ. Trims are obtained by dropping a fraction of its digits, say, those indexed by multiples of 2^k: dropping every second digit, or every fourth, eighth, etc. Call such models internal if γ is generic itself.

Internal models respect a family **IM** of axioms that do eliminate all non-integer quantifiers. These axioms assert, for almost every sequence γ, the equivalence of each closed formula F to its modification F' that replaces all 2nd order variables α_i by algorithms $A_i(\gamma^{(k_i)})$ computing those sequences from trims of that γ, with k_i and algorithms A_i quantified as integers. *Cf.* [4].

3 Algorithmic Information: A Brief Background

3.1 Basic Concepts

I freely use concepts involving quantifiers only over hereditarily finite sets (I often call "integers"). Presuming reader knows how to develop such basics as, e.g., algorithms, I do not try to work out elegant ways to define them directly in membership terms. Below $\|x\| \stackrel{\rm df}{=} n$ for $x{\in}\{0,1\}^n$ and $\|t\| \stackrel{\rm df}{=} \lceil \log_2 t \rceil{-}1$ for $t{\in}\overline{\mathbb{R}^+}$. ***Uniform measure*** on $\mathbf{\Omega} \stackrel{\rm df}{=} \{0,1\}^{\mathbb{N}}$ (or on $\mathbf{\Omega}^k \simeq \mathbf{\Omega}$) is $\lambda(x\,\mathbf{\Omega}) \stackrel{\rm df}{=} 2^{-\|x\|}$.

Partial continuous transforms (PCT) on $\mathbf{\Omega}$ may fail to narrow-down the output to a single sequence, leaving a compact set of eligible results. So, their graphs are compact sets $A \subset \mathbf{\Omega}{\times}\mathbf{\Omega}$ with $A(\alpha) \stackrel{\rm df}{=} \{\beta : (\alpha,\beta){\in}A\} \neq \emptyset$. Singleton outputs $\{\beta\}$ are interpreted as $\beta{\in}\,\mathbf{\Omega}$.

Preimages $A^{-1}(s) \stackrel{\rm df}{=} \{\alpha : A(\alpha) \subset s\}$ of open $s \subset \mathbf{\Omega}$ are always open.

Closed A also have closed preimages of all closed s.

Computable PCTs have algorithms enumerating the clopen subsets of $\mathbf{\Omega}^2 \backslash A$. $U(p\alpha)$ is a universal PCT. It computes $n{=}\|x\|$ bits of $A_p(\alpha){\subset}x\mathbf{\Omega}$ in $\mathbf{t}_{p\alpha}(n)$ steps (and $U(p){\in}\mathbb{N}$ in $\mathbf{t}_p$ steps).

A ***computably enumerable*** (c.e.) function to $\overline{\mathbb{R}^+}$ is sup of a c.e. set of basic continuous ones.

Dominant in a Banach space C of functions is its c.e. $f\in C$ if all c.e. g in C are $O(f)$. Such is $\sum_i g_i/(i^2{+}i)$ if (g_i) is a c.e. family of all c.e. functions in the unit ball of C.

λ-test is $\mathbf{d}(\alpha)\stackrel{\rm df}{=}\|\lceil\mathbf{T}(\alpha)\rceil\|$ for a c.e. $\mathbf{T}$: $\mathbf{\Omega}{\to}\overline{\mathbb{R}^+}$, $\lambda(\mathbf{T}){\le}1$, dominant in $\mathbf{L}^1(\mathbf{\Omega},\lambda)$.

Kolmogorov–Martin-Lof (K-ML) λ-random are $\alpha\in\mathbf{R}^\lambda_\infty$, where $\mathbf{R}^\lambda_c\stackrel{\rm df}{=}\{\alpha:\mathbf{d}(\alpha)<c\}$ (compact for $c\in\mathbb{N}$). Let $\mathbf{M}(x)\stackrel{\rm df}{=}\lambda(U^{-1}(x\mathbf{\Omega}))$. $\mathbf{R}^\lambda\stackrel{\rm df}{=}\mathbf{R}^\lambda_\infty$ consists of all γ with $\sup_{x:\,\gamma\in x\mathbf{\Omega}}\frac{\mathbf{M}(x)}{\lambda(x\,\mathbf{\Omega})}<\infty$.

Mutual Information (dependence) $\mathbf{I}(\alpha_1:\alpha_2)$ is $\min_{\beta_1,\beta_2}\{\mathbf{d}((\beta_1,\beta_2)) : U(\beta_i){=}\alpha_i\}$.

Measuring classes. For a sequence of clopen $F_i{\subset}\mathbf{\Omega}$ with $\mathbf{M}(F_i{\oplus}F_{i+1}){<}2^{-i}$, let open $F^+_i\stackrel{\rm df}{=}\cup_{j>i}F_j$, compact $F^-_i\stackrel{\rm df}{=}\cap_{j>i}F_j$; $F^-\stackrel{\rm df}{=}\cup_i F^-_i{\subset}F^+\stackrel{\rm df}{=}\cap_i F^+_i$. Any $F{\in}\Delta^0_*$ has such $(F_i){\in}\Delta^0_*$ with $F^-{\subset}F{\subset}F^+$. Note that $F^+\backslash F^-$ contains no *generic* γ. If $F\subset\mathbf{R}^\lambda$ and $\lambda(F)=0$ then also $\mathbf{M}(F_i)=O(2^{-i})$, implying $\mathbf{I}(\gamma:f)=\infty$ for $f{=}(F_i)_i$ and all $\gamma{\in}F$. Thus, any arithmetic class F includes no generic $\gamma\in\mathbf{R}^\lambda$ if $\lambda(F){=}0$, and includes all such γ or none if F is invariant under single digit flips.

3.2 Weak Truth-table (Closed PCT) Reductions to Generic Sequences

$\mathbf{P}\boldsymbol{\chi}$ Turing-reduces each $\beta\in\mathbf{\Omega}$ to a $\alpha\in\mathbf{R}^\lambda$: $\beta=U(\alpha)$. PCT U can use unlimited segments α_m, discarding nearly all their information, to produce small segments β_n. But **IP** makes $\mathbf{P}\boldsymbol{\chi}$ equivalent to its stronger form, requiring a closed PCT u, with $m\sim n$. Yet, by [9], one cannot require a total u, nor α Turing-equivalent to β: some information loss cannot be avoided.

Let $s^n_t\stackrel{\rm df}{=}\lambda(\{\alpha{:}\ \mathbf{t}_\alpha(n){<}t\})$, $s_\infty\stackrel{\rm df}{=}\inf_n s^n_\infty$. The computable function τ^n_r is $\min\{t : s^n_t{>}r\}$, $r{\in}\mathbb{Q}$.

Let α be generic. Then $s_\alpha\stackrel{\rm df}{=}\liminf_n s^n_{\mathbf{t}_\alpha(n)}<s_\infty$ as $\lambda(\{\alpha : s_\alpha{=}s_\infty\})=0$. Take $r\in(s_\alpha,s_\infty)$ and a monotone infinite sequence n^α_i of all n with $\mathbf{t}_\alpha(n){<}\tau^n_r$. Then $\forall^\infty i\ \min\{p:\mathbf{t}_p>n^\alpha_i\}<i\|i\|^2$.

$U_c(\alpha){\in}\{\#,0,1\}^{\mathbb{N}}$ avoids divergence by diluting $U(\alpha)$ with $\min_{p<i\|i\|^2+c}\{\mathbf{t}_p{:}\ \mathbf{t}_p{>}n^\alpha_{i+1}\}\le\infty$ blanks # after n^α_i-th bits. U_c carries no extra information of α absent in $U(\alpha),n^\alpha$.
As U_c never diverges, $\mu\stackrel{\rm df}{=}U_c(\lambda)$ is computable and $U'(\alpha)\stackrel{\rm df}{=}\mu([\#^{\mathbb{N}},U_c(\alpha)])$ maps $\mathbf{R}^\lambda$ to $\mathbf{R}^\lambda\cup U(\mathbb{N})$.

From $\beta=U'(\alpha){\in}\mathbf{R}^\lambda$ we recover $U_c(\alpha)$ and $u(\beta)\stackrel{\rm df}{=}U(\alpha)$. PCT u is closed (a w.t.t. reduction) as the input segments it uses are only slightly longer (by codes p for n^α bounds) than the output's.

(Viewing $\#^+\{0,1\}^+$ segments of $U_c(\alpha)$ as integers makes μ a computable (on $\mathbb{N}^{<\mathbb{N}}$ prefixes) distribution on (finite and infinite) sequences of integers. U' gives them short codes.)

3.3 To Eliminate 2nd Order Quantifiers with IP

Elimination of non-integer quantifiers in predicates with free variables meets an obstacle in collision-resistant one-way functions over $\mathbf{\Omega}$: computable, injective only on sets of measure 0, and yet allowing only a 0 chance of generating "siblings" – distinct inputs with the same outputs. See [1].

Thus our task of reducing quantifiers to integer ones is restricted to only closed sentences, with no free variables. The Internal Models axioms (**IM** at the end of Sec.2) achieve that. Yet, adding the whole **IM** family strikes me as less elegant or intuitive than a single neat axiom, such as $\mathbf{P}\boldsymbol{\chi}$.

Open Problems and Further Research. But **IM** may be just theorems. Any statement G consistent with **IP**, $\mathbf{P}\boldsymbol{\chi}$ holds in a countable model M of the axioms. Let such G be $F\,\&\,\widetilde{F}$ for $\widetilde{F}\stackrel{\rm df}{=}\lambda(\{\gamma{:}\ F'(\gamma)\}){=}0$. Then M for G respects $\mathbf{P}\boldsymbol{\chi}$ and **IP** with some $\Psi\in\Delta^0_*\cap\mathbf{R}^\lambda$ such that $\forall\gamma\,((\widetilde{F}\,\&\,F'(\gamma))\Rightarrow\gamma\notin\mathbf{R}^\lambda_\Psi)$. where $\mathbf{R}^\lambda_\Psi\stackrel{\rm df}{=}\{\gamma:(\Psi,\gamma){\in}\mathbf{R}^\lambda_c\}$ for some fixed c. For a contradiction we need to express M in an "internal style", enumerating its sets from trims $\gamma^{(k)}$ of $\gamma\in\mathbf{R}^\lambda_\Psi$. If a reduction basis of M is Turing-equivalent to $\alpha_i{\in}\mathbf{R}^\lambda_\Psi$, we need to merge α_i's into such γ as its trims.

Would be great to so prove **IM** as theorems. If not, they may still be equivalent to a single axiom with a clear common-sense intuition or at least to a single axiom whose intuition may need more research to be fully understood. If this fails, too, one may still try to replace **IM** with some family offering a more compelling intuition, of the sort available for, e.g., **IP**.

4 Some Discussion

Cantor Axioms license on formula-defined sets led to fatal consistency problems. Zermelo-Fraenkel's (somewhat *ad-hoc*) cardinality-based restrictions diffused those but left intact their self-referential root. Seems just an example of uncoiling a vicious circle into a vicious spiral. The result was a Babel Tower of cardinalities, other hierarchies finding little relevance in math. And the generality of this height of hierarchies is illusory. Expanding Set Theory with more formula types, axioms, etc. has no natural limit. Benefits are few and eventual consistency loss inevitable.

A clean way out may be recognizing the distinction between collections internal to math, specified by its formulas, and external ones that math handles as values of variables, without specifying.

Internal collections have a limited hierarchy: the type of allowed formulas is clear-cut. Any extension would make a new theory, with its own clear limits. External objects would be fully independent of internal ones: having finite information about them. Complexity theory allows to formalize that, justify the validity for "external data," and use that for simplifying math foundations.

General math objects are collections specified by formulas with external sets as parameters. Collections of them are treated as collections of those parameters. So, any such object would rely on a single formula, not on all of them, thus excluding back-door extensions of formula language.

The uniform set concept of all math objects with no explicit types hierarchy is luring but illusive. The hierarchy of ever more powerful axioms, models, cardinals remains, if swept under the rug. Making it explicit and matching the math relevance may be a path to simpler foundations.

What is left out? – "Logical" sets, related to infinite hierarchies of formulas, such as "the set of all true sentences of Arithmetic." Those should be subject of math foundations. Theories cannot include their own foundations. Math Logic then could focus on math rather than on itself.

Acknowledgment. I am grateful to Albert Meyer for helpful criticism.

References

[1] George Barmpalias, Xiaoyan Zhang. 2025. Collision-resistant hash-shuffles on the reals. https://arxiv.org/abs/2501.02604

[2] H.Friedman. 1974. Some systems of second-order arithmetic and their use. *ICM proc.* **1**:235-242.

[3] Randall Holmes. 2006. Alternative Set Theories. *Stanford Encyclopedia of Philosophy.* https://plato.stanford.edu/entries/settheory-alternative/

[4] Leonid A. Levin. 1976. On the Principle of Conservation of Information in Intuitionistic Mathematics. *Soviet Math. Dokl.* **17**/2:601-605 = *DAN SSSR* **227**/6:1293-1296.

[5] Leonid A. Levin. 2013. Forbidden Information. *JACM*, **60/2**. Sec. "The Independence Postulate." https://arxiv.org/abs/cs/0203029

[6] Leonid A. Levin. 2022. Gacs-Kucera Theorem. *TCS*, **929**:172-173. https://arxiv.org/abs/2207.12489

[7] Stephen G. Simpson. 2009. *Subsystems of Second Order Arithmetic.* Cambridge U. Press.

[8] E.Zermelo. 1908. Untersuchungen über die Grundlagen der Mengenlehre. *Math.Ann.* **65**:261-281.

[9] Vladimir V. V'yugin. The Algebra of Invariant Properties of Binary Sequences. *Probl.Inf.Transm.* 18/2:147-161, 1982.

A Appendix. Some Other IP Applications

Foundations of Probability. Paradoxes in its application led to the K-ML randomness concept $\mathbf{R}^\lambda$. **IP** clarifies its use: For any $S\subset\mathbf{\Omega}$: $\lambda(S)=0$ if and only if $\exists\sigma\ S\cap\mathbf{R}^\lambda\subset\{\gamma:\mathbf{I}(\gamma:\sigma)=\infty\}$.

Goedel Theorem Loophole. Goedel writes:

> "It is not at all excluded by the negative results mentioned earlier that nevertheless every clearly posed mathematical yes-or-no question is solvable in this way. For it is just this becoming evident of more and more new axioms on the basis of the meaning of the primitive notions that a machine cannot imitate."

No way ! Let a predicate P on $\{0,1\}^n$ extend "proven/refuted" partial predicate of Peano Arithmetic. Let r_n be the n-bit prefix of a c.e. real $r = \min \mathbf{R}_0^\lambda$. Then $\mathbf{I}(P : r_{[1,n]}) = n \pm O(\log n)$. No way to obtain such P for any significant n, by formal or informal methods !

B Appendix. ZFC Axioms

ZFC axioms are sometimes given to undergraduates in an unintuitive, hard to remember list. Setting them in three pairs seems to help intuition.

Sets with a given Set Theory property F (possibly with parameters c) are said to form a ***class*** $\{x : F_c(x)\}$. They may or may not form a set, but only sets are the domain of ZFC variables.

1. **Membership chains: sources, sinks**. (1b anti-dual to 1a):
   - **1a. Infinity** (a set with no membership source): $\boxed{\exists S{\neq}\emptyset\, \forall x{\in}S \exists y{\in}S (x{\in}y)}$
   - **1b. Foundation** (each set has sinks: members disjoint with it): $\boxed{\neg \exists S{\neq}\emptyset\, \forall x{\in}S \exists y{\in}S (y{\in}x)}$
2. **Sets with formula-defined membership:**
   - **2a. Extensionality:** (content identifies sets uniquely): $\boxed{x{\supset}y{\supset}x{\in}t \Rightarrow y{\in}t}$
   - **2b. Replacement:**[2] $\boxed{(\forall x\, \exists Y{\supset} R_c(\{x\})) \Rightarrow \forall X\, \exists Y{\supset} R_c(X){\supset}Y}$
3. **Functions Inverses**. $f^{-1} \overset{\text{df}}{=} \{g \subset f^T : f(g(f(x))) = f(x)\}$:
   - **3a. Power Set** ($f^{-1} \subset G$ is a set, take $h = f^T$): $\boxed{\forall h \exists G\, \forall g{\subset}h\, (g{\in}G)}$
   - **3b. Choice**[3] (f^{-1} is not empty): $\boxed{\forall f\, \exists g \in f^{-1}}$

The modifications discussed above include:

1. Restrict Replacement to computable R. Add its opposite, the Independence Postulate.
2. The Power Set, too, is replaced with its opposite: "All sets are countable."
3. Add the Primal Chaos axiom. (It is an open question if something stronger is needed. At worst, it could be the Internal Model axioms family.)
4. Extend Foundation and Extensionality axioms to classes, as axioms families;
5. Some change with the Choice. (May be dropping it, or adding to the language a postulated (named, not described) class that injects reals into countable v.Neumann ordinals (continuum hypothesis implied).

Math objects are classes $\{q$: $F_p(q)\}$ of sets q satisfying formulas F with external parameters p. Collections of objects are treated as collections of those parameters. Quantifiers bind parameters, not properties F.

Papers may have families of theorems parameterized by formulas. Such are used now for the so called Large Categories that are not Zermelo-Fraenkel sets. They are also used for families of axioms, like Induction axioms. We may treat such families as meta-mathematical statements. Alternatively, one may allow variables for formulas but not quantifiers on them, except for an implicit universal quantifier over all free variables in a sentence. So parameterized will be the Foundation and Extensionality axioms, because they must apply to classes, not just sets.

[2] An axiom for each Set-Theory-defined relation $R_c(x,y)$. $R_c(X) \overset{\text{df}}{=} \{y : \exists x{\in}X\, R_c(x,y)\}$.
[3] The feasibility of computing inverses is the most dramatic open problem in Computer Theory.