Zermelo-Fraenkel Axioms, Internal Classes, 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. They have finite "depth": are described by a single Set Theory formula with variables ranging over objects unrelated to math formulas. Exotic expressions referring to sets with no depth limit or to Powerset axiom 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, treated as the domain of variables) to be hereditarily countable and independent of formula-defined classes, i.e. with finite Kolmogorov Information about them. This allows elimination of all non-integer quantifiers in Set Theory statements.

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 Set Theory (ST) - 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 are pure ***classes*** defined by math formulas. The other are ***externals*** that math handles but generally does not specify internally. Random sequences are an example.[1] Externals are proper sets and are the domain of ST variables. I postulate them to be hereditarily countable and independent of pure classes, i.e. have only finite Kolmogorov Information about them. **Math objects** (only informally called sets) are classes $\{q\colon F_p(q)\}$ of sets q satisfying formulas F with external parameters p. Collections and properties of objects are treated as ones of those parameters. Excluded would be "Logics" objects of infinite "depth" 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 must preserve it; only a separate Power Set increases it.

This cardinality focus has problematic relevance. Distinctions between uncountable cardinalities almost never looked at in math papers. Besides, generic sets based on Power Set Axiom, with no other descriptions, find little use in math and greatly complicate its foundations.

And as all axiom systems have countable models, cardinalities feel like artifacts, designed to hide self-referential aspects. Many papers in Logic (e.g., [1, 2, 6]) 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: a very unfortunate effect. And complicating proofs is unattractive, too.

Expanding Set Theory with fancier formula types, axioms, etc. has no natural end. Benefits are little, and eventual consistency loss inevitable. Let us try to drop any such excesses.

Some Formalities. In effect, I split math objects into a hierarchy of ranks, each quantifier binding only classes below one rank. Rank $\le k$ classes F^k_p are specified by a hereditarily countable ("proper" set) parameter p via the universal Σ^1_k formula $F^k_p(q)$. In p, a v.Neumann ordinal $o(p)$ is included. $F^k_p(q)$ defines membership $F^k_q{\in}F^k_p$ conditioned on $o(q){<}o(p)$. Membership is also extended by extensionality: $F^{k'}_{q'} \in F^k_p$ if $F^{k'}_{q'} = F^k_q$, i.e. have a formula-defined isomorphism of membership relations on their transitive closures. Typical math concepts have straightforward translations. For instance, "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.)

*Boston University, CAS, Computer Science department, 665 Commonwealth Ave., 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 great simplifying approximation to negligibly small ϵ, so is ∞ for $\frac{1}{\epsilon}$. And note that the set $\overline{\mathbb{R}}$ of infinitely long reals is compact, "less infinite" in that than $\mathbb{Q}$.

1 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 ST formulas $F(n)$ 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.

Proper sets are external data which math handles without specifying internally. Sets constitute the domain of ST variables. They can be put in collections based on properties and relations with other sets. This forms "mixed" classes, specified by ST formulas with some free variables taken for data parameters. Data may be chaotic, but "simpleminded," playing no self-referential tricks.

General math objects – mixed classes – carry both tricks and chaos. But their tricks are limited to those of a single formula. Distinguishing sets from pure classes allows another radical insight: Even with infinite complexity, external objects $\alpha \in \mathbf{\Omega} \stackrel{\text{df}}{=} \{0,1\}^{\mathbb{N}}$ can have only finite dependence on pure sequences Ψ. Complexity theory (see sec. 2 for some background) allows to formalize that as having finite (small, really) mutual Kolmogorov Information $\mathbf{I}(\alpha : \Psi) < \infty$. Note, recursive and random transformations can only add finite information about a given (e.g., by a formula) sequence.

[4] discusses the validity of this Independence Postulate (**IP**) for external data. (Besides, it is redundant for math objects F_p to duplicate formula-defined information in external parameters p.) Mixed math objects would not encroach on infinite hierarchies. They reduce to a pair of sequences: one in $\Delta^0_* \stackrel{\text{df}}{=} \cup_n \Delta^0_n$ and another that carries no information about pure classes. This gives a way to handle infinitely complex sets, but reduce their quantifiers to those on **integers**. All with no seeming need to change anything in math papers, only reinterpret some formalities.[2]

By **IP**, only computable sequences can double as both pure classes and sets. So, this family of axioms conflicts with Replacement axioms for sets. (For classes these axioms are merely definitions.) Thus they must be restricted to just one: "Being sets is preserved by Turing reductions." Instead, the Foundation axiom becomes a family, like Induction axioms: one for each class of ordinals.

And "Primal Chaos" axiom ($\mathbf{P}\boldsymbol{\mathcal{X}}$) is added: "Each $\alpha \in \mathbf{\Omega}$ is computable from [3] even-indexed digits of some Martin-Lof random β: $\mathbf{d}(\beta) < \infty$." (For classes, i.e. in ZFC, it is a theorem: see [5].)

Remarkably, **IP** opens a way to eliminate all non-integer quantifiers, It excludes α satisfying $F(\alpha, \beta)$ unless such α reduce to a positive fraction of those $\gamma \in \mathbf{\Omega}$ from which β is computable, too. The converse requires some version of $\mathbf{P}\boldsymbol{\mathcal{X}}$. Here is a (countable) model (in ZFC):

Generic are $\gamma \in \mathbf{\Omega}$ outside all arithmetic sets $X \in \Delta^0_*$ of measure $\lambda(X)=0$. Take one. γ-model includes all sets with "$\in$" on transitive closures enumerable from $\gamma^{(k)}$ for some k. (In this model, 2nd order quantifiers are eliminated with a c.e. family of axioms: $F \Leftrightarrow \lambda(\{\gamma \colon F'(\gamma)\}) > 0$ where F' has all real variables α_i in a sentence F replaced with $A_i(\gamma^{(k_i)})$; A_i, k_i quantified as integers. *Cf.* [3].)

2 Kolmogorov Information: A Brief Background

Uniform measure on $\mathbf{\Omega}$ (or on $\mathbf{\Omega}^2 \simeq \mathbf{\Omega}$) is $\lambda(x\,\mathbf{\Omega}) \stackrel{\text{df}}{=} 2^{-\|x\|}$. (For $x \in \{0,1\}^n$, $\|x\| \stackrel{\text{df}}{=} n$.)

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 $A(\alpha) \stackrel{\text{df}}{=} \{\beta : (\alpha, \beta) \in A\} \neq \emptyset$. Singleton outputs $\{\beta\}$ are interpreted as $\beta \in \mathbf{\Omega}$.

Preimages $A^{-1}(s) \stackrel{\text{df}}{=} \{\alpha : A(\alpha) \subset s\}$ of all open sets $s \subset \mathbf{\Omega}$ in any PCT A are open.

Closed A also have closed preimages of all closed s. For $x \in \{0,1\}^*$ let $x^A \stackrel{\text{df}}{=} \{\alpha : A(\alpha) \in x\mathbf{\Omega}\}$.

---

[2] Some statements can be a meta-theorem: a family with formula parameter F, as done now by Category Theorists.
[3] "Computable from" here can be equivalently restricted to "weak truth-table reducible to" (see sec. 2.1).

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 $t_{p\alpha}(n)$ steps (and $U(p)\in\mathbb{N}$ in 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.

$\boldsymbol{\lambda}$-test is $\mathbf{d}(\alpha)\stackrel{\rm df}{=}\|\lceil\mathbf{T}(\alpha)\rceil\|$ for a dominant $\mathbf{T}$ on $\mathbf{\Omega}$ with expectation $\lambda(\mathbf{T}){\le}1$. $\|t\|\stackrel{\rm df}{=}\lceil\log_2 t\rceil{-}1$.

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

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 $\mathbf{M}(F_i{\oplus}F_{i+j}){<}2^{-i}$, 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 $F\subset\mathbf{R}^\lambda$, $\lambda(F)=0$ then also $\mathbf{M}(F_i)<2^{-i}$, implying $\mathbf{I}(\gamma:(F_i)_i)=\infty$ for 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.

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

Let $s^n_t\stackrel{\rm df}{=}\lambda(\{\alpha\colon t_\alpha(n){<}t\})$, $s_\infty\stackrel{\rm df}{=}\inf_n s^n_\infty$. The function $\tau^n_r\stackrel{\rm df}{=}\min\{t:s^n_t{>}r\}$, $r{\in}\mathbb{Q}$, is computable. Let $\mathbf{I}(\alpha{:}\,s_\infty){<}\infty$. Then $s_\alpha\stackrel{\rm df}{=}\liminf_n s^n_{t_\alpha(n)}<s_\infty$ as $s_\infty{\notin}\Delta^0_2$. Take $r{\in}(s_\alpha,s_\infty)$ and a monotone infinite sequence n^α_i of all n with $t_\alpha(n){<}\tau^n_r$. Then $\forall^\infty i\ \min\{p:t_p{>}n^\alpha_i\}<i\|i\|^2$.

$U_c(\alpha){\in}\{\#,0,1\}^{\mathbb{N}}$ avoids divergence by diluting $U(\alpha)$ with $\min_{p<i\|i\|^2+c}\{t_p\colon t_p{>}n^\alpha_{i+1}\}{\le}\infty$ 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 $\mathbb{N}^*$ prefixes) distribution on (finite and infinite) sequences of integers. U' gives them short codes.)

2.2 Elimination of 2nd Order Quantifiers with IP

Simplifications with $\mathbf{P}\boldsymbol{\mathcal{X}}$. To ***reduce*** all ST predicates or classes, i.e., express them using only integer quantifiers, we need to reduce $\{\beta:\exists\alpha\,(\alpha,\beta){\in}P\}$, for any reduced P, where $\beta\in\mathbf{\Omega}^k{\simeq}\mathbf{\Omega}$. Note, $\exists\alpha{\in}\mathbf{\Omega}\,P(\alpha)\Leftrightarrow\exists\alpha{\in}\mathbf{R}^\lambda_0\,P(u(\alpha))$, so α,β,γ below will range only over $\mathbf{R}^\lambda$. Also, any pair (α,β) is $(A(\gamma),u(\gamma))$, so we need to reduce $u(P')$ for reduced $P'\stackrel{\rm df}{=}\{\gamma\colon\exists A(A(\gamma),u(\gamma)){\in}P\}$.

Let $\lambda_x(Q)\stackrel{\rm df}{=}\frac{\lambda(Q\cap x^u)}{\lambda(x^u)}$, $\lambda_\beta(Q)\stackrel{\rm df}{=}\liminf_{x\sqsubset\beta}\lambda_x(Q)$ be λ on γ conditioned on $u(\gamma){=}\beta$. Note: $\lambda_\beta(Q){>}t$ is reduced if Q is. **IP** assures $\lambda_\beta(P'){>}0$ for $\beta{\in}u(P')$, so taking $P_k\stackrel{\rm df}{=}\ (P'\cup\{\gamma\colon\lambda_{u(\gamma)}(P'){\le}2^{-k}\})^-$ turns $\beta{\in}u(P')$ into $\exists k\,(\beta{\in}u(P_k)\&\,\lambda_\beta(P'){>}2^{-k})$. So we are to reduce $u(P_k)$ or $F\stackrel{\rm df}{=}u(P_k){\cap}\mathbf{R}^\lambda$. Note, $\forall x\,\lambda_x(P_k)>2^{-k}$, so it seems $F=\mathbf{R}^\lambda$. But proving that is complicated by one-way functions issue:

Recursively one-way functions on S, are closed computable PCT A, s.t. $\lambda(x^A\cap S)=O(\lambda(x\mathbf{\Omega}))$, $\lambda(S){>}0$, and no computable B invert A, i.e. $\lambda^2(\{(\beta,\gamma)\colon\alpha\stackrel{\rm df}{=}B(\beta,\gamma){\in}S,A(\alpha){=}\beta\})=0$.

If no such OWF existed, $\mathbf{P}\boldsymbol{\mathcal{X}}$ would give a reverse of the above exclusion by **IP**, thus eliminate all non-integer quantifiers. But handling OWFs demands a stronger than $\mathbf{P}\boldsymbol{\mathcal{X}}$ addition to **IP**. A model described above achieves that with a c.e. family of axioms. Yet, adding such a family does not strike me as really elegant and intuitive. A single axiom strengthening $\mathbf{P}\boldsymbol{\mathcal{X}}$ would be neater. Or expanding **IP** can be tried. $\mathbf{R}^\lambda\backslash F$ has $<1{-}2^{-k}\ u(\lambda)$-density in each $x\mathbf{\Omega}$. Thus $\mathbf{M}(\mathbf{R}^\lambda\backslash F)=0$ and $\mathbf{I}(\alpha{:}\beta){=}\infty$ for for some β and all $\alpha{\in}\mathbf{R}^\lambda\backslash F$. But F has one $\mathbf{\Omega}$-ranged quantifier. So **IP** would need an extention beyound Δ^0_* for that use. I am not sure yet, which solution would work best.

3 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. The result was a Babel Tower of cardinalities, other hierarchies finding little relevance in math. A clean way out may be recognizing the distinction between internal to math collections 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.

Generic 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 swiped 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.

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

References

[1] H.Friedman. 1974. Some systems of second-order arithmetic and their use. *ICM proc.* **1**:235-242.
[2] Randall Holmes. 2006. Alternative Set Theories. *Stanford Encyclopedia of Philosophy.* https://plato.stanford.edu/entries/settheory-alternative/
[3] 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.
[4] Leonid A. Levin. 2013. Forbidden Information. *JACM*, **60/2**. Sec. "The Independence Postulate." https://arxiv.org/abs/cs/0203029
[5] Leonid A. Levin. 2022. Gacs-Kucera Theorem. *TCS*, **929**:172-173. https://arxiv.org/abs/2207.12489
[6] Stephen G. Simpson. 2009. *Subsystems of Second Order Arithmetic.* Cambridge U. Press.
[7] E.Zermelo. 1908. Untersuchungen ber die Grundlagen der Mengenlehre. *Math.Ann.* **65**:261-281.

A Appendix. Some of 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 $r=\min\mathbf{R}^\lambda_0$. Then $\mathbf{I}(P:r_n)=n\pm O(\log n)$.

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 ST 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:[4] $\boxed{(\forall x\, \exists Y{\supset}\, R_c(\{x\})) \;\Rightarrow\; \forall X\, \exists Y{\supset}R_c(X){\supset}Y}$

3. **Functions Inverses**. $f^{-1} \stackrel{\text{df}}{=} \{g\colon f(g(x)){=}x,\ \text{Dom}(g){=}\text{Im}(f)\}$:

   3a. Powerset ($f^{-1} \subset G$ is a set, as $g \subset h = f^T$): $\boxed{\forall h \exists G\, \forall g\, (g{\subset}h \Rightarrow g{\in}G)}$

   3b. Choice[5] (f^{-1} is not empty): $\boxed{\forall f\, \exists\, g \in f^{-1}}$

The modifications discussed above include:

1. Restrict Replacement to computable R, drop Power Set;
2. add **IP**, **P$\boldsymbol{\chi}$**, strengthened;
3. extend Foundation to classes of sets (as a family of axioms, like Peano Induction);
4. some (unclear yet) replacement for Choice.[6]
   (May be dropping it, or adding to the language a postulated (not described) class mapping countable v.Neumann ordinals onto reals, implying continuum hypothesis, too.)

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

[4] An axiom for each ST-defined relation $R_c(x, y)$. $R_c(X) \stackrel{\text{df}}{=} \{y : \exists x{\in}X\, R_c(x, y)\}$.

[5] The feasibility of computing inverses is the most dramatic open Computer Theory problem.

[6] I am not clear yet on how to handle some math theorems that depend on Axiom of Choice, as it provides classes not specified by ZF formulas. One option may be extending the language with a postulated (not otherwise described) class that maps countable v.Neumann ordinals onto reals (implying Continuum Hypothesis, too).
But any such options would require careful analysis.