Independence, randomness and the axiom of choice

van Lambalgen, M.

Published in:
Journal of Symbolic Logic

DOI:
10.2307/2275368

THE JOURNAL OF SYMBOLIC LOGIC
Volume 57, Number 4, Dec. 1992

INDEPENDENCE, RANDOMNESS AND THE AXIOM OF CHOICE

MICHIEL VAN LAMBALGEN

Abstract. We investigate various ways of introducing axioms for randomness in set theory. The results show that these axioms, when added to ZF, imply the failure of AC. But the axiom of extensionality plays an essential role in the derivation, and a deeper analysis may ultimately show that randomness is incompatible with extensionality.

§0. Introduction. The fact that ZF set theory does not decide various easily expressible statements concerning the reals and sets of reals leads one to search for additional axioms. Most axioms proposed so far use concepts which are formalizable in set theory and they express new, in some cases even plausible, properties of these concepts. Typical examples include large cardinal axioms, Martin's axiom, or the axiom of determinacy. But, as Kreisel pointed out, one might also think of a different way of introducing axioms, this time involving new primitives:

> Let us try to expand the language of set theory, that is, add symbols for new primitive notions and look for axioms in the wider language, which are evident (for the notions given). They may imply set theoretic propositions, i.e., propositions in the language L_E of set theory, which are not. Prima facie the case for this project is overwhelming. [Kreisel then goes on to discuss one such proposal, namely, adding a truth predicate.]
>
> A more interesting, but also more problematic, expansion in the literature [Kruse 1967] concerns the primitive predicate of being a "*random sequence.*" Myhill formulated axioms for the notion l.c., which, however, are not altogether plausible. As he himself observed, the *intuitive* notion of random sequence is certainly *not* extensional; specifically, if we think of a sequence as being produced by a die, it is random, but this does not ensure that some element of the sequence differs from, say, 1; also, it is not intuitively clear that there is a collection of (intensional) sequences such that (i) for each sequence it is decided whether or not the sequence is random and (ii) the usual axioms of set theory are satisfied by the collection. So, once again, we have here an idea for an expansion rather than an effective use of an expansion ([1968], 100–1).

In philosophical discussions of the cumulative hierarchy it is quite common to introduce additional primitives to capture the underlying picture. For example,

Received October 11, 1990; revised December 18, 1991.

0022-4812/92/5704-0007/$04.10

Boolos [1985] added the primitives "stage" and "earlier than," and Van Aken [1986] added "presupposes." But it will be seen below that these new concepts by no means capture all there is to say about the cumulative hierarchy. In particular, consideration of a predicate for randomness is suggested by Gödel's remarks on CH and ordinal definability. In *The consistency of the axiom of choice and the generalized continuum hypothesis* we read

> The proposition [V = L] added as a new axiom seems to give a natural completion of the axioms of set theory, in so far as it determines the vague notion of an arbitrary infinite set in a definite way ([1990], 27).

Of course, Gödel later revised his opinion about the plausibility of V = L, but he never doubted that the "vague notion of an arbitrary infinite set" would have to be determined in some way in order to decide, e.g., the continuum hypothesis:

> If, however, someone believes that it is meaningless to speak of sets except in the sense of extensions of definable properties, or at least, that no other sets exist, then, too, he can hardly expect more than a small fraction of the problems of set theory to be solvable without making use of this, in his opinion essential characteristic of sets, namely, that they are extensions of definable properties. This characteristic of sets, however, is neither formulated explicitly nor contained implicitly in the accepted axioms of set theory (*What is Cantor's continuum problem*? [1990], 183).

In 1946, he proposed ordinal definability as a way to make the latter characteristic precise:

> I would think that "definability in terms of ordinals" [...] is at least an adequate formulation in an absolute sense for [the] property "being formed according to a law" as opposed to "being formed by a random choice of the elements." For in the ordinals there is certainly no element of randomness, and hence neither in sets defined in terms of them. This is particularly clear if you consider von Neumann's definition of ordinals, because it is not based on any well-ordering relation of sets, which may very well involve some random element (*Remarks before the Princeton bicentennial conference* [1990], 152).*

These quotations immediately suggest the question: is it possible (and desirable) to give a parallel characterization of "being formed by a random choice of the elements"? Mutatis mutandis, Gödel's remarks about definability apply here: if one believes that it is essential that one can form *arbitrary* (not just definable) subsets

*There is, as far as I know, one other place where Gödel alludes, somewhat engimatically, to a connection between randomness and set theory. In a letter to Tarski he writes, commenting on the failure of his "square axioms" for the continuum: "My confidence that $2^{\aleph_0} = \aleph_2$ has of course somewhat been shaken. But it still seems plausible to me. One reason is that I don't believe in any kind of irrationality such as, e.g. random sequences in any absolute sense." (Gödel [1990], 175)

of a given set, then the axioms should reflect this. In the set theoretical literature, it is customary to interpret "arbitrary" by "as many as possible," which means that one tries to use some kind of maximum principle to motivate additions to ZF. However, the results of applying a maximum principle are not unambiguous; in particular, one can use such principles to argue both for and against CH (see Maddy [1988]). The argument *against* CH is familiar: why would $\mathbb{R}$, which is obtained by a "bold new operation" (Cohen's phrase), have the same cardinality as the set of countable ordinals, which is merely the simplest way of generating an uncountable set? Consider, however, the following shrewd argument *for* CH (due to John Steel): $\mathbb{R}$ and a well-ordering of length ω_1 are created at stage $\omega + 2$, a bijection between $\mathbb{R}$ and a well-ordering of length ω_1 would appear at stage $\omega + 3$, and it would artificially restrict the power set operation at *that* stage if we did not include the bijection.

A different interpretation of "arbitrary" was suggested by Paul Bernays (and before him by Borel in [1909]), in terms of the mode of generation of subsets. He wrote in *On Platonism in mathematics*

> ... we can imagine functions engendered by an infinity of independent determinations which assign to each integer an integer, and we reason about the totality of these functions. In the same way, one views a set of integers as the result of infinitely many independent acts deciding for each number whether it should be included or excluded. We add to this the idea of the totality of these sets. ([1985], 260)

In this paper, we will explore the mathematical consequences of Bernays's point of view, and the notion of independence will be fundamental here. First a remark on terminology. Henceforth we shall call a sequence obtained by infinitely many independent choices a *random* sequence. (This is slightly at variance with the usual terminology, where randomness also refers to statistical properties such as stability of relative frequencies.) At first sight, it may seem impossible to give a mathematical characterization of a random sequence as "the result of infinitely many independent acts." A priori, a sequence obtained in this way will not satisfy any "special" property, i.e., a property which does not follow from the sequence being a real number; and in classical logic we are then forced to say that the sequence does not have any special properties at all. Below it will be shown that this contradiction can be avoided if we consider absolute independence to be an ideal, which can only be approached by means of a suitable hierarchy, much like the universe V is approached by the levels V_α. In fact, our main axiom will be a kind of reflection principle for this hierarchy. A difference with the cumulative hierarchy is that we have to formulate the new hierarchy in terms of a new primitive called *independence*. The axioms for the hierarchy imply that there is no well-ordering of the reals (hence the failure of AC), and conversely, if there is no well-ordering of the reals, then the axioms for the hierarchy can be interpreted (the latter statement will be made precise in §1). In other words, one can also extract *a proof* of the existence of a well-ordering of the reals from these pages. In this context it should be mentioned that Bernays obviously thought that his informal use of the notion of independence is compatible with the axiom of choice; he even claims (without proof) that AC is evident in this picture. This poses a challenge to those who believe in choice: can

one set up nontrivial axioms for independence, which are compatible with, or even imply, the axiom of choice? Actually, in the above quotation Bernays seems to present *two* notions of independence, one for functions and one for sets. The latter can be reduced to the first, but not conversely. If we consider the independence notion for functions to be fundamental, then AC is not implausible; but if one tries to imagine an informal proof of AC in the case where the independence notion for sets is taken to be fundamental (as we do here), one will see that AC is concerned rather with *dependence* of successive choices.

Let us now proceed to a brief description of the contents of this paper. In §1 we give the axioms together with philosophical motivation and proof of consistency, we show by a simple diagonal argument that AC fails in this context and we state some results about conservative extensions. In §2 we study the axioms by means of a generalized quantifier, H. Friedman's "almost all." This quantifier allows us to prove the conservation results mentioned in §1 in a compact way; but the results of this section can be interpreted alternatively as an approach to randomness, which axiomatizes only quantification over random sequences, instead of introducing a predicate for (and hence a set of) random sequences. Nevertheless, it will be proved here that these approaches are essentially equivalent, despite their philosophical differences.

In §3 we take up the matter of extensionality. In ordinary set theory, sets are thought of as "being simply there," we are not at all concerned with the way sets are presented to us, hence the axiom of extensionality is obvious. On this viewpoint, random sequences are extensionally different from nonrandom sequences. On the other hand, one might argue that randomness concerns the mode in which sequences are given to us, and in this case it is more reasonable to give up extensionality. We indicate here why the diagonal argument against AC does not go through in the absence of extensionality. (Of course one does not really salvage AC in this way, because in the absence of extensionality AC loses much of its usual force.) In §4 we introduce some further plausible axioms for randomness, which are analogous to the density and data axioms for the (intuitionistic) theory of lawless sequences. A consequence of the new axioms is the existence of a translation invariant extension of Lebesgue measure to all sets of reals. To appreciate the potential significance of this result, recall that, in modern probability theory as codified by Kolmogorov, one proceeds by setting up first the apparatus of σ-algebras and σ-additive measures; afterwards the notions of independence and randomness are defined. The results of this article indicate that one can profitably reverse this procedure: we start with the notions of randomness and independence as true primitives; it is then shown that one can dispense with the usual apparatus.

In §5 we briefly investigate the relation between the set theories studied in the previous sections and intuitionistic set theory extended with lawless sequences. Lawless sequences provide another interpretation of "arbitrary": to generate a real, the ideal mathematician freely chooses one integer after another. (Referring back to the quotation from Bernays: we have a sequence of "independent acts," but we do not add the idea of "the totality of these sets.") In this article, we will repeatedly turn to this concept as a source of inspiration, and actually the classical and intuitionistic notions of "arbitrary" can be related formally by means of an

interpretation of set theory with independence axioms into intuitionistic set theory with lawless sequences.

In an appendix, §6, we compare our approach to Freiling's "Axioms of Symmetry" [1986]. Freiling shows how some simple axioms on throwing n random darts forces the cardinality of the continuum to be $\geq \aleph_n$. The main result here is, roughly, that the size of $2^{\aleph_0}$ is related to the number of iterations of Q that one allows. Freiling himself draws the conclusion that the cardinality of the continuum must be at least as large as $\aleph_{\omega+1}$. An examination of his proofs shows that he really establishes the nonexistence of well-orderings of the continuum of length ω_n. It will be clear from the above that we conclude from this that it is rather AC, which is doubtful.

In this section we also comment on the axioms for randomness proposed by Myhill, which were referred to in the quotation from Kreisel.

It will have become clear by now that the main issues in this investigation are (1) should we conceive of randomness as intensional or extensional?, (2) is randomness compatible with classical logic? (Prima facie, random sequences are incomplete objects; can we conceive of them as being completed?), and (3) is randomness compatible with the usual axioms of set theory?

With regard to the main issues singled out above, our conclusions are that (1) randomness can be treated both extensionally and intensionally, (2) we have found no properties that force us to give up classical logic, and (3) one has to give up either extensionality or choice. It seems to us that the simplicity of the proof of 1.3 (failure of AC), combined with the generality of the axioms for independence, show that these axioms describe a nonartificial notion of set for which AC is false; roughly speaking, they describe a universe in which we do not have perfect information about all sets. The other possibility (failure of extensionality) also appears to be interesting, but so far we have not been able to provide a consistency proof for ZFC-EXT plus the randomness axioms.

Acknowledgment. I would like to thank Domenico Zambella for permission to include some of his results and Chris Freiling, Georg Kreisel and Penelope Maddy for helpful correspondence. The author is supported by a Huygens-fellowship of the Netherlands Organization for Scientific Research (NWO).

§1. An extensional study of randomness. The task now at hand is to describe axiomatically some important properties of sequences of independently chosen integers, that is, of random sequences. One intuition behind randomness is, that we cannot say a priori that a random sequence will satisfy any given nontrivial property. The main difficulty here arises from the use of classical logic. For the intuitionistic theory of lawless sequences embodies the aforementioned intuition exactly: we imagine a mathematician freely choosing integers, such that his choices are not subject to any restriction; at any particular moment our information concerning a lawless sequence consists of a finite initial segment only, and excluded middle is not valid for properties depending on an infinite number of coordinates (e.g., the property that all coordinates are zero). Classical logic forces us to say that if we do not want to assume that a sequence satisfies a certain nontrivial property, then we must assume that it does not satisfy this property, and hence the informal notion of random sequence seems to be inconsistent.

One way out of this difficulty consists in combining a relativized form of randomness with the use of what might be called a reflection principle. The ordinary reflection principle tries to express the immensity of the universe V by emphasizing the poverty of the language used to describe V: any property true of V is already true of a set (usually a level V_α of the cumulative hierarchy). Iteration of this principle in fact shows that this property must then hold for arbitrarily large V_α. In other words, there exist arbitrarily large set-like approximations of the universe. We will try to express the richness of the power set in the same way: we introduce levels of randomness (this corresponds to the levels of the cumulative hierarchy), and we require that if a property holds for a random x of a certain level, it also holds for random x of higher levels (this corresponds to the reflection principle). The skeptical reader may object that there is not much of an analogy here, since language is too poor to describe V (which exists) because it is too large, whereas language cannot describe absolute randomness because it is inconsistent, i.e., does not exist. Our answer to this is that in the second case we also stumble over the poverty of language: because a theory formulated in classical logic cannot make as much distinctions as one formulated in intuitionistic logic, we are forced to conclude that absolute randomness is inconsistent. In intuitionistic logic, however, this concept makes sense, and the axioms given below "reflect" properties of the intuitionistic concept in the classical universe, as the interpretation defined in §5 will make clear. It should be observed here that there is also a formal analogy between the two reflection principles, in that both correspond to generalized quantifiers, which can be interpreted as "almost all": for the ordinary reflection principle it is Shelah's *aa* (for applications in set theory, see, e.g., Kaufmann [1983] and Kakuda [1989]), in our case it is Friedman's "almost all," which will be investigated in the next section.

We therefore introduce a relativized randomness notion $R(x, \vec{y})$ with the intended interpretation "x is random relative to $\vec{y}$". Here, x ranges over infinite binary sequences and $\vec{y}$ ranges over finite tuples of arbitrary sets. ($\vec{y}$ denotes a vector, *of unspecified length*, of variables; hence R is a relation of indefinite arity.)

If $R(x, \vec{y})$, we also say that *x is independent of $\vec{y}$* or that *$\vec{y}$ has no information about x*. $\vec{y}$ may be empty, in which case we write $R(x, \varnothing) =: R(x)$. $R(x, \varnothing)$ really means that the "definable" sets contain no information about x. The phrase "has no information" should be understood in the following way (cf. also 2.9): the language at our disposal does not allow us to define (with parameters $\vec{y}$) a "small" set in which x is contained. (So that the axioms given below emphasize the limitations of our language and of our thinking.) This use of the term "information" was suggested by the analogy with Kolmogorov complexity developed in van Lambalgen [1990].

We now present the axioms for the independence relation; comments will be given afterwards. Properties 1–4 and 6 are a slight weakening of the usual axioms for algebraic or linear independence. Observe that R6 corresponds to the Steinitz exchange principle. It would be of interest to investigate whether a category theoretic definition of independence (see, e.g., Cohn [1981], 254) would yield a more elegant approach, but we have not pursued this.

R0. Axioms and inference rules for classical predicate logic

R1. $\exists x R(x)$,

R2. $R(x, \vec{y}\vec{z}) \to R(x, \vec{z})$,
R3. (a) $R(x, \vec{y}) \to R(x, \pi\vec{y})$ for any permutation π; (b) $R(x, y\vec{z}) \to R(x, yy\vec{z})$,
R4. $R(x, y) \to x \neq y$.

The next axiom is the reflection scheme:

R5. Suppose $\phi(x, \vec{y})$ is in $\{\in, =, R\}$, where x ranges over reals and all parameters are listed among the $\vec{y}$. Then

$$\exists x(R(x, \vec{y}) \wedge \phi(x, \vec{y})) \to \exists x(R(x, z\vec{y}) \wedge \phi(x, \vec{y})).$$

(Of course this statement is nontrivial only if z does not occur free in ϕ.)

R6. $R(y, \vec{z}) \wedge R(x, y\vec{z}) \to R(y, x\vec{z})$.

Occasionally we shall refer to another axiom, which determines the behaviour of R with respect to ordinals:

R7. Let α be an ordinal. Then $R(x, \vec{y}) \to R(x, \alpha\vec{y})$.

The system consisting of axioms R0–R6 will be denoted $\mathscr{R}$. If we want to include R7, we indicate this as $\mathscr{R}(1-7)$.

Comments on the axioms. (1) This expresses that there is a sequence, which is random relative to a given universe of "definable" sets.

(2) This axiom introduces a partial order among the levels of randomness. In fact, R determines a labelled (or decorated) lattice: the elements of the lattice are given by the $\vec{y}$ ($\varnothing$ is the bottom element), the ordering by inclusion, and join and meet by union and intersection respectively; and to each element $\vec{y}$ we attach the label $\{x \mid R(x, \vec{y})\}$. In this way, one can view R as defining a frame for modal predicate logic: the nodes correspond to worlds, the labels to domains. The only unusual feature is that in this case the domains are nonincreasing, whereas ordinarily they are nondecreasing. The levels have been defined only for finite sequences of parameters $\vec{y}$, but the results given below can be used to show that this hierarchy can be extended to well-ordered sequences of parameters. Note that one cannot say that $R(x, \vec{y}\vec{z})$ determines a higher level of randomness than $R(x, \vec{z})$; e.g., if the $\vec{y}$ are ordinals, the levels can be taken to be the same (cf. R7).

It will be observed that 2 would follow from the transitivity of $\neg R$, a property that is usually included in axiomatizations of independence. Since this property would preclude closure of random sequences under nontrivial operations such as subsequence selection, we cannot use it. (The interpretation into the theory of lawless sequences shows that one can consistently add transitivity; however, the universe of lawless sequences is notoriously not closed under nontrivial continuous operations.) Perhaps it is of interest to observe that, under the correspondence between generalized quantifier properties and independence properties defined in the proof of 2.6(i), transitivity corresponds to "a countable union of countable sets is countable," i.e., the characteristic axiom for the quantifier "there are uncountably many..." (see van Lambalgen [1991]).

(3) This axiom, which says that the argument $\vec{y} = y_1 \cdots y_n$ should really be thought of as the set $\{y_1, \ldots, y_n\}$, is added only for convenience, because it facilitates the formulation of the other axioms.

(4) It seems obvious that if x is independent of y, or if y carries no information about x, then x must in some sense be different from y. But do we mean different *extensionally* (*qua* sets) or *intensionally* (*qua* names of sets)? Are random sequences

extensionally different from "definable" sequences, or are they simply given in a different way? In the presence of the axiom of extensionality we have no choice, but we might want to give up extensionality to accommodate these distinctions; this possibility will be investigated in §3.

(5) The intuitive justification of this axiom runs as follows. Suppose $\phi(x, \vec{y})$ is satisfied by some random x of level $\vec{y}$. Like the universe V, absolute randomness cannot be characterized in our language. However, we would like to *think* of this particular x as being absolutely random, and one way to express this is that ϕ is satisfied by random x of arbitrary high level.

The reader may wonder why we have to start at level $\vec{y}$, where the $\vec{y}$ comprise all the parameters (except x) of ϕ. This is to avoid inconsistency. Otherwise we could have, e.g., $\exists x(R(x) \wedge x = y) \rightarrow \exists x(R(x, y) \wedge x = y)$, whence we get by contraposition, $\forall x(R(x, y) \rightarrow x \neq y) \rightarrow \forall x(R(x) \rightarrow x \neq y)$. By R4 we get $\forall x(R(x) \rightarrow x \neq y)$, hence R must be empty, in contradiction with R1.

R5 has the following consequence: if f is a function definable in $\{\in, =, R\}$, then $\forall x(R(x, \vec{y}) \rightarrow x \neq f(\vec{y}))$. For suppose that $\exists x(R(x, \vec{y}) \wedge x = f(\vec{y}))$. Put $z := f(\vec{y})$, then R5: $\exists x(R(x, z\vec{y}) \wedge x = f(\vec{y}))$, which conflicts with R4. In particular we have that x satisfying $R(x)$ will not be definable in $\{\in, =, R\}$. The same type of reasoning shows that $R(x, \vec{y}) \rightarrow R(x, f(\vec{y}))$. For suppose that for some $\vec{y}$, $\exists x(R(x, \vec{y}) \wedge \neg R(x, f(\vec{y}))$. If we put for this particular $\vec{y}$, $z := f(\vec{y})$, then we have by R5: $\exists x(R(x, z\vec{y}) \wedge \neg R(x, f(\vec{y}))$, which contradicts R2. Hence x satisfying $R(x)$ are also *independent* from all definable sets.

Another consequence of R5 is a strengthening of R1: $\forall \vec{y} \exists x R(x, \vec{y})$, by means of the following chain of inferences: $\exists x R(x)$ (R1), $\exists x(R(x, y) \wedge x = x)$ (from R5), etc. In particular we have that for some x, $R(x, 2^\omega)$. This may strike the reader as strange, but it means that we cannot extract much information from 2^ω using the language $\{\in, =, R\}$ with parameter 2^ω. It will be obvious that, in this picture, 2^ω has no definable well-ordering.

Further clarification of R5, in terms of its relation to forcing, will be given after the proof of Lemma 4.4.

(6) The Steinitz exchange principle is harder to motivate intuitively. We think of random sequences as being indistinguishable: after all, since random sequences do not have special properties, there can be no property that distinguishes one from the other. Roughly speaking, this would mean that if we take two different random sequences, then there should be a permutation of the structure, which exchanges the two. This intuition cannot be fully reproduced in a classical framework (although it holds in models for lawless sequences) and R6 is the classical approximation to our intuition.

(7) We shall see later that R6 allows us to prove a form of Gödel's idea that "in the ordinals there is certainly no element of randomness." In our set-up this idea corresponds to R7: if a real is independent of a set, it is independent of that set *cum* ordinal, i.e., ordinals do not count. R6 is mainly responsible for the fact that the addition of R7 to ZF$\mathscr{R}$ is conservative for formulas in the language of set theory (see 2.9). Once we have R7, we can also assume that we need not mention ordinals in the R predicates occurring in R5. Hence the results mentioned in (5) also hold for ordinal definable functions. An easy consequence is that there is no ordinal

definable function f from an ordinal onto the reals; for by what was said under (5), $\forall\alpha\forall x(R(x,\alpha) \to x \neq f(\alpha))$ and since we may drop the ordinal in R we get $\forall x(R(x) \to \forall\alpha(x \neq f(\alpha)))$, so that by R1 the range of f cannot be the set of all reals.

Let us first collect the properties of R mentioned so far.

1.1. LEMMA. (i) ($\mathscr{R}$). *If f is definable in $\{\in, =, R\}$, then $\forall x(R(x,\vec{y}) \to x \neq f(\vec{y}))$ and $\forall x(R(x,\vec{y}) \to R(x, f(\vec{y})))$*

(ii) ($\mathscr{R}(1-7)$). *If f is ordinal definable, then $\forall x(R(x,\vec{y}) \to x \neq f(\vec{y}))$ and $\forall x(R(x,\vec{y}) \to R(x, f(\vec{y})))$*

(iii) $\forall\vec{y}\exists x R(x,\vec{y})$.

We now study the effect of adding $\mathscr{R}$ to ZF. For the purpose of consistency proofs, it is convenient to study two possibilities: the parameters $\vec{y}$ can be interpreted by arbitrary sets, or they range over reals only. The system ZF$\mathscr{R}$ results from ZF by adding $\mathscr{R}$, where $\vec{y}$ can be interpreted by arbitrary sets. The system ZF$\mathscr{R}^0$ results from ZF by adding $\mathscr{R}$, such that the $\vec{y}$ range over reals. (The philosophical motivation for the second system is somewhat different from that of the first: it tries to come close to the intuitionistic idea that two different sequences of independent choices must themselves be independent.) In both cases we have the following side conditions

(1) R may occur in the schemata,

(2) the first argument of R is assumed to range over 2^{ω}.

Here are some easy consistency results:

1.2. LEMMA. (i) *ZF$\mathscr{R}$ minus R5 is conservative over ZF; similarly for ZF$\mathscr{R}^0$*

(ii) *if T is a first-order theory with an infinite model, then $T + \mathscr{R}$ is conservative over T; in particular, if we do not allow R in the schemata of ZF then ZF $+ \mathscr{R}$ is conservative over ZF.*

PROOF. (i) If $\vec{y}$ denotes the vector $\langle y_1, \ldots, y_n\rangle$, $\#(x,\vec{y})$ is defined as: $x \neq y_1 \wedge \cdots \wedge x \neq y_n$. Now interpret R as: $R(x,\vec{y}) \Leftrightarrow \#(x,\vec{y})$.

(ii) See van Lambalgen [1990], Theorem 2.1.3. □

We shall show presently that ZF$\mathscr{R}$ falsifies the axiom of choice.

1.3. THEOREM (ZF$\mathscr{R}$). *There is no choice function on the power set of the reals.*

PROOF. Suppose g were such a choice function. Then $\{x \mid R(x,g)\}$ is a set by the separation axiom; moreover, it is nonempty in virtue of R1 and R5 (cf. Lemma 1.1). Hence we can apply g to this set, that is, we have $R(g\{x \mid R(x,g)\}, g)$, or in other words, $\exists x(R(x,g) \wedge x = g\{x \mid R(x,g)\})$.

By R5 we must have, for any z, $\exists x(R(x,zg) \wedge x = g\{x \mid R(x,g)\})$. If we put $z := g\{x \mid R(x,g)\})$, we obtain a contradiction with R4. □

The essence of the proof is contained in Lemma 1.1(i) and (iii): that if x is independent of y, then x cannot be the value of a function F applied to y; for y, take the choice function g and let F be the ($\{\in, =, R\}$-definable) function that associates to each choice function g' on the power set of the reals, the real $g'\{x \mid R(x,g')\}$. 1.1(i) seems to be a reasonable requirement for a notion of independence, so the burden of the proof is carried by the transcendentality property 1.1(iii). If one is willing to accept $\forall\vec{y}\exists x R(x,\vec{y})$ only for real $\vec{y}$, one can still refute the existence of a projective well-ordering. Note that, even when the question of the plausibility of ZF$\mathscr{R}$ is set aside, the construction of Theorem 1.3 gives us some information about what goes on in universes of set theory (like the world of AD) in which all sets of

reals are Lebesgue measurable, or have the Baire property, because in those cases (as we shall see later) we can introduce an independence relation obeying R1–R7.

The following partial converse to 1.3 will be proved in the next section. Let AC_{wo} denote the statement that all well-ordered collections of nonempty sets have a choice function.

1.4. THEOREM. *There exists an interpretation of* $ZF\mathscr{R}(1-5,7)$ *into* $ZF + AC_{wo} +$ "*there is no well-ordering of the reals*".

In other words, one can derive the existence of a well-ordering of the reals if one assumes that there is no "no information predicate" described by 1–5, 7. Or, to put it in the form of a slogan: "omniscience implies well-ordering of the reals." Is there a connection with intuitionistic set theory, where AC implies the law of excluded middle? (The use of extensionality is essential here!; see Diaconescu [1975] and Goodman and Myhill [1978].)

We now turn to matters of consistency. We first give a description of the (well-known) model that we shall use. Let $\mathscr{M}$ be a countable transitive model of ZF + V = L. In $\mathscr{M}$, consider $(2^\omega)^\kappa$, where $\kappa \geq \omega_1$. We equip $(2^\omega)^\kappa$ with the product topology and the product measure λ^κ defined on the Borel σ-algebra $B((2^\omega)^\kappa)$. Let I denote the σ-ideal of λ^κ nullsets, then the quotient algebra $\mathscr{B} := B((2^\omega)^\kappa)/I$ is a complete Boolean algebra. Let $\mathscr{G} \subseteq \mathscr{B}$ be a generic ultrafilter; construct the generic extension $\mathscr{M}[\mathscr{G}]$. We shall refer to this extension as "generically adding κ random reals."

For any sequence $\vec{y}$ of elements in $2^\omega \cap \mathscr{M}[\mathscr{G}]$, $\mathscr{M}[\vec{y}]$ is well-defined (e.g., via relative constructibility).

In $\mathscr{M}[\mathscr{G}]$, interpret $R(x, \vec{y})$ as

$$R(x, \vec{y}) \text{ iff for all Borel sets } B \text{ with code in } \mathscr{M}[\vec{y}], \text{ if } \lambda B = 1, \text{ then } x \in B.$$

If $R(x, \vec{y})$, we say that x is (Solovay) random over $\mathscr{M}[\vec{y}]$.

1.5. LEMMA. $R(x, \vec{y})$, *interpreted as Solovay randomness, satsifies* $\mathscr{R}^0$.

PROOF. R1 holds because the set of reals random over $\mathscr{M}[\vec{y}]$ has outer measure 1 in $\mathscr{M}[\mathscr{G}]$. R2–R4 are trivial and R6 was verified in van Lambalgen [1990], Theorem 2.2.1. To prove R5, we first observe that R is in fact definable; hence it suffices to show that for ϕ in the language of ZF (we consider the contraposition of R5):

$$\mathscr{M}[\mathscr{G}] \models \forall x(R(x, z\vec{y}) \to \phi(x, \vec{y})) \to \forall x(R(x, \vec{y}) \to \phi(x, \vec{y})).$$

Suppose $\mathscr{M}[\mathscr{G}] \models \forall x(R(x, z\vec{y}) \to \phi(x, \vec{y}))$. If x is an element of $\mathscr{M}[\mathscr{G}]$, let $\underline{x}$ denote its name. We claim that there exists a formula ψ such that $\mathscr{M}[\mathscr{G}] \models \phi(x, \vec{y}) \Leftrightarrow \mathscr{M}[\vec{y}, x] \models \psi(x, \vec{y})$. By the homogeneity of $\mathscr{B}$, we can write $\mathscr{M}[\mathscr{G}] = \mathscr{M}[\vec{y}, x][\mathscr{G}']$, where $\mathscr{G}' \subseteq \mathscr{B}$ is a generic ultrafilter. If π is any automorphism on $\mathscr{B}$ such that the induced automorphism on $\mathscr{M}[\mathscr{G}]$ fixes $\vec{y}$ and x, then $\mathscr{M}[\vec{y}, x][\pi\mathscr{G}'] \models \phi(x, \vec{y})$. It follows that $\mathscr{M}[\vec{y}, x] \models$ "$\mathbf{1}_{\mathscr{B}} \Vdash \phi(\underline{x}, \underline{\vec{y}})$" so that "$\mathbf{1}_{\mathscr{B}} \Vdash \phi(\underline{x}, \underline{\vec{y}})$" is the required formula $\psi(x, \vec{y})$. By the properties of Solovay forcing, there exists a Borel set B, with code in $\mathscr{M}[\vec{y}]$, such that for all x random over $\mathscr{M}[\vec{y}]$, $x \in B \Leftrightarrow \mathscr{M}[\vec{y}, x] \models \psi(x, \vec{y})$.

By hypothesis, for all x random over $\mathscr{M}[\vec{y}, z]$, $\mathscr{M}[\vec{y}, x] \models \psi(x, \vec{y})$. Since this set has outer measure 1, we must have $\lambda B = 1$. But then, because B has code in

$\mathcal{M}[\vec{y}]$, also for all x random over $\mathcal{M}[\vec{y}]$, $\mathcal{M}[\vec{y}, x] \models \psi(x, \vec{y})$. This shows $\mathcal{M}[\mathcal{G}] \models \forall x(R(x, \vec{y}) \to \phi(x, \vec{y}))$. □

1.6. Theorem. (a) ZF$\mathcal{R}^0$ + AC *can be interpreted in the model obtained by generically adding κ random reals, for $\kappa \geq \omega_1$.*

(b) ZF$\mathcal{R}$ + DC *is consistent.*

Proof. (a) Interpret R as Solovay randomness and apply 1.5. Since R is definable, the separation and substitution schemata with the new predicate R are also satisfied.

(b) Consider the submodel HOD(2^ω). □

1.7. On the Steinitz exchange principle. The reader will have observed that so far we have not used R6, which expresses an indistinguishability property of random sequences. In the next section it will become clear that R6 plays an important role in the proof that ZF$\mathcal{R}$(1–7) is conservative (in $\{\in, =\}$) over various subsystems. We mention the most important examples here. The meaning of the first result has already been explained.

1.7.1. Theorem. ZF$\mathcal{R}$(1–7) *is* $\{\in, =\}$*-conservative over* ZF$\mathcal{R}$.

For the next result we need a

1.7.2. Definition. The class I of *independence formulas* is the smallest class of formulas in the language $\{\in, =, R\}$ such that

(i) if ϕ is a formula in $\{\in, =, R\}$, then ϕ is in I;

(ii) I is closed under $\wedge, \neg, \to, \forall$;

(iii) if $\phi(x, \vec{y})$ is in I (all free variables are indicated), then $\forall x(R(x, \vec{z}) \to \phi(x, \vec{y}))$ is also in I, where $\vec{z}$ contains $\vec{y}$.

In other words, in formulas of the class I, the R predicate is used only to relativize quantifiers. Let us denote by ZF$\mathcal{R}(I)$ the system that results from ZF$\mathcal{R}$ when we restrict the formulas ϕ in R5 to I formulas, and when only I formulas are allowed in the schemata. The philosophical motivation behind ZF$\mathcal{R}(I)$ is the following. It might be objected to ZF$\mathcal{R}$ that it requires the collections $\{x \mid R(x, \vec{z})\}$ to be sets. In the weaker system this is no longer required: here we can quantify over sequences random at a certain level, but the domain of quantification need not be a set. (Again, note the analogy between V and R.) We then have

1.7.3. Theorem. ZF$\mathcal{R}$ is $\{\in, =\}$*-conservative over ZF$\mathcal{R}(I)$.*

1.8. Remark. Clearly the notion of "arbitrary subset of..." is not confined to subsets of the natural numbers, whereas the axiom system $\mathcal{R}$ deals only with these. One would like to have an R predicate on all power sets, not just the reals; in particular it would be more elegant to allow also iteration of R, as in formulas of the form: $R(\{x \mid R(x)\})$. Perhaps some form of Easton forcing can be used to prove consistency.

§2. Quantification over random sequences. The results of this section can be viewed in two ways. The first is to see the introduction of the quantifier "almost all" as a technical device to make explicit the notion of largeness implicit in the reflection principle R5. In the case of the ordinary reflection principle, the notion of largeness is given by a quantifier for closed unbounded sets of ordinals (see Kaufmann [1983] and Kakuda [1989]). It will be seen below that in our case the associated quantifier resembles Lebesgue measure.

The second way is to view the quantifier in its own right as a formalization of some aspects of randomness. When we generate a random sequence, there is prima facie nothing certain about the result, except for the finite segment already generated. In this sense, random sequences behave as lawless sequences and are indeed often adduced as examples of the latter. However, on an intuitive level we are also almost certain, or practically certain, of many other properties: that relative frequencies converge, that the same holds in suitably chosen subsequences, etc. The fact that practical certainty is not absolute certainty is what renders an extensional theory of randomness difficult. But we have a different option: formalizing the properties of "practically certain" itself and incorporating this notion in ZF.

So suppose we have some stochastic mechanism X that randomly produces infinite binary sequences x. For X one could take, e.g., a fair coin. We add to the language of set theory a generalized quantifier Q, where the intended interpretation of $Qx\phi(x)$ is: "if x is randomly generated, it is practically certain that $\phi(x)$," or more concisely: "almost certainly, ϕ." We shall state the axioms first, and comment on them afterwards. Unless specified otherwise, we assume that the variables bound by Q run over 2^ω. We first fix the logic to be used:

Q0. Axioms and inference rules of classical predicate logic.

The following five properties, expressing that Q is a nonprincipal filter, are fairly immediate.

Q1. $\neg Qx\, x \neq x$,

Q2. $Qx\, x \neq y$,

Q3. $Qx\phi(\ldots,x,\ldots) \rightarrow Qy\phi(\ldots,y,\ldots)$ provided y is free for x in ϕ,

Q4. $Qx\phi \wedge \forall x(\phi \rightarrow \psi) \rightarrow Qx\psi$,

Q5. $Qx\phi \wedge Qx\psi \rightarrow Qx(\phi \wedge \psi)$.

The crucial property is

Q6. $QxQy\phi \leftrightarrow QyQx\phi$.

Comments on the axioms. Q0. This choice of a logic is not altogether obvious, since we have expressly introduced random sequences as objects about which we have partial information only, and in general such objects are governed by intuitionistic logic.

Q1. When read classically, this axiom states that there is some randomly generated x. Without such a weak existence axiom, the enterprise would trivialize. But notice that Q1 is still *very* weak; the following density axiom also seems to be justified: if a set A has positive measure, then $\neg Qx(x \notin A)$. This stronger axiom will be considered in §4.

Q2. This means that it is almost certain that randomly generated sequences differ from some given sequence. This axiom is related to Kreisel's observation that "if we think of a sequence as being produced by a die, it is random, but this does not ensure that some element of the sequence differs from, say, 1"; for y, take the sequence 111.... In this way we bring in an element of intensionality, without giving up the axiom of extensionality: we keep extensional $=$, but we only require that it is *almost* certain that a randomly produced x differs from y.

Q3. This is a syntactic expedient only.

Q5. States that the intersection of two almost certain events is again almost certain (though, intuitively, perhaps slightly less so).

Q4. Expresses that Q is a monotone quantifier. That this is less innocent than may seem at first sight can be seen as follows. We show that $Qx\phi(x) \rightarrow \exists x\phi(x)$. Suppose $Qx\phi(x)$; assume for simplicity that ϕ does not contain other free variables. Since $\forall x(\phi(x) \rightarrow \exists y\phi(y))$, monotonicity implies $Qx\exists y\phi(y)$. Using classical logic and Q1–Q5 one can derive, when x does not occur free in $\exists y\phi(y)$: $Qx\exists y\phi(y) \rightarrow \exists y\phi(y)$. For suppose $\neg\exists y\phi(y)$, then also $\forall x(x \neq y \rightarrow \neg\exists y\phi(y))$, whence by Q4 and Q2, $Qx\neg\exists y\phi(y)$. By Q5, $Qx(\neg\exists y\phi(y) \wedge \exists y\phi(y))$, which, again by Q4, contradicts Q1. Hence $\neg\neg\exists y\phi(y)$, and so $\exists y\phi(y)$ by classical logic. To see why the existential import of Q is somewhat problematic, let us make a historical diversion. When Borel had proved (in [1909]) that almost all real numbers are absolutely normal (i.e., normal in every base), he noted that it would be of interest either to construct a concrete example of such a number or to show that none of the definable numbers is absolutely normal. He observed that, however paradoxical at first sight, the latter possibility did not contradict his result. In other words, the existential quantifier in $Qx\phi(x) \rightarrow \exists x\phi(x)$ need not have a constructive meaning. It should be observed here that the addition of Q1–Q5 to ZF is conservative, so that no new existential statements are provable by means of $Qx\phi(x) \rightarrow \exists x\phi(x)$. This situation changes, however, as soon as one adds Q6.

Q6. We think of the quantifiers Qx, Qy in a formula $QxQy\phi$ as referring to independent processes. Hence $QxQy\phi(x, y)$ means: "if x is randomly generated and y is randomly generated independently from x, then it is practically certain that $\phi(x, y)$." But if the processes generating x and y are independent, it should not matter in which order we take them. (This intuition is also the motivation behind Freiling's "Axioms of Symmetry" [1986].)

The axiom system Q0–Q6 will be denoted $\mathscr{Q}$. It will be observed that we can think of the formula $Qx\phi(x)$ as meaning "$\{x \mid \phi(x)\}$ has Lebesgue measure 1." On this interpretation, Q6 bears some analogy to the Fubini theorem: for Fubini implies Q6 *if we know in addition that* $\{\langle x, y\rangle \mid \phi(x, y)\}$ *is measurable*. Without this extra condition, Q6 expresses a very strong property, which is responsible for the peculiar features of set theory with random sequences. Despite the fact that there is only an analogy, we shall often refer to Q6 as the "Fubini property." The measure theoretic interpretation was studied by Harvey Friedman, who in fact was the first to introduce a quantifier with properties 1–6 (see Steinhorn [1985a, b]). This interpretation motivates occasional references to sets A such that $Qx(x \notin A)$ as "small sets" or "nullsets."

We now investigate the consequences of adding $\mathscr{Q}$ to ZF; after that, we turn to matters of consistency. We first show how *not* to add $\mathscr{Q}$ to ZF. Generally, we have

2.1. LEMMA. *Let T be any first-order theory. Then $T + \mathscr{Q}$ is conservative over T.*

PROOF. See van Lambalgen [1990], Theorem 2.1.3. □

In particular, ZFC + $\mathscr{Q}$ is conservative over ZFC, even when the variables bound by Q run over *all* sets. However, the theory ZFC + $\mathscr{Q}$ is unnatural, since we do not allow Q to occur in the schemata of ZF; that is, once we have decided that the notion of randomness is important enough to consider in the context of set theory, we might as well go the whole way and allow sets defined in terms of it. We therefore consider only extensions of set theory in which Q does occur in the separation and substitution schemata. Again, it is useful to distinguish two exten-

sions of this sort: ZF$\mathscr{Q}$ and ZF$\mathscr{Q}^0$. In ZF$\mathscr{Q}$, Q formulas may contain parameters for arbitrary sets; in ZF$\mathscr{Q}^0$, only real and ordinal parameters are allowed in Q formulas. Hence in ZF$\mathscr{Q}$ we allow formulas "$Qx(x \in A)$" for *arbitrary* sets A in 2^ω; in the weaker theory ZF$\mathscr{Q}^0$ we still have formulas "$Qx(x \in B)$", with B a *projective* subset of 2^ω.

Prima facie it seems that systems of this type (unlike the ones studied in §1) are fairly weak in one respect: one cannot express that random sequences are closed under nontrivial operations. Nevertheless, even in the context of Zermelo set theory, the axioms pose strong constraints on the set theoretic universe. We first give a simple result, which shows that axioms Q1–Q5 are fairly innocent.

2.2. LEMMA. *ZF$\mathscr{Q}$ minus Q6 is conservative over* ZF.

PROOF. Define $Qx\phi$ explicitly as "$\{x \mid \phi(x)\}$ is cofinite," then Q1–Q5 are satisfied. □

On the other hand, if we add Q6, we get the following strong additivity property of Q:

2.3. THEOREM (ZF$\mathscr{Q}$). *Let the family $(A_\alpha)_{\alpha<\kappa}$ be given, where κ is an ordinal. Then $\forall\alpha < \kappa Qx(x \notin A_\alpha) \to Qx\forall\alpha < \kappa(x \notin A_\alpha)$.*

PROOF. (The argument is adapted from van Benthem [1990]). We use induction on ordinals κ. Note that we need Q in the separation axiom to conclude from this that the result holds for *all* families $(A_\alpha)_{\alpha<\kappa}$. We may suppose that A_α are pairwise disjoint. The case $\kappa = 1$ is trivial. So suppose for all $\lambda < \kappa$ and all families $(B_\alpha)_{\alpha<\lambda}$, $\forall\alpha < \lambda Qx(x \notin B_\alpha) \to Qx\forall\alpha < \lambda(x \notin B_\alpha)$. Define a relation S on 2^ω by $S(x, y) \Leftrightarrow \forall\alpha < \kappa[x \in A_\alpha \to (\exists\beta > \alpha(y \in A_\beta) \vee \forall\beta < \kappa(y \notin A_\beta))] \vee [\forall\alpha < \kappa(x \notin A_\alpha) \wedge y \in 2^\omega]$.

Set $B := \bigcup_{\alpha<\kappa} A_\alpha$. We then have $\forall xQyS(x, y)$. For if $x \notin B$, then $\{y \mid S(x, y)\} = 2^\omega$; and if $x \in B$, then $\{y \mid S(x, y)\} = 2^\omega - C$, where C is a union of less than κ A_α's; now apply the induction hypothesis. Hence also $QxQyS(x, y)$, whence by Q6, $QyQxS(x, y)$. Consider the set $\{x \mid S(x, y)\}$. Obviously, if $y \in B$, then $\{x \mid S(x, y)\} = B^c \cup C$, where C is a union of less than κ A_α's; and if $y \notin B$, then $\{x \mid S(x, y)\} = 2^\omega$. It follows from this observation that $\{y \mid QxS(x, y)\} \supseteq B^c$. If we have equality, then $QxQyS(x, y)$ implies $Qx(x \notin B)$. If not, then for some $y \in B$, $QxS(x, y)$; but $\{x \mid S(x, y)\} = B^c \cup C$, where C is a union of less than κ A_α's, and we can apply the induction hypothesis to obtain $Qx(x \notin C)$. But $Qx(x \in B^c \vee x \in C) \wedge Qx(x \notin C)$ implies $Qx(x \in B^c)$. □

In particular, we see that the continuum is not the union of κ Q-nullsets, for *any* ordinal κ.

2.4. COROLLARY (ZF$\mathscr{Q}$). *Suppose $B \subseteq 2^\omega$ has a well-ordering. Then $Qx(x \notin B)$.*

PROOF. Suppose $B = \{x_\alpha \mid \alpha < \kappa\}$. By Q2, $Qx(x \notin \{x_\alpha\})$. Now apply the preceding theorem. □

2.5. COROLLARY. (ZF$\mathscr{Q}$) $\neg AC$.

PROOF. If not, 2^ω would have a well-ordering, whence by the preceding corollary, $Qx(x \notin 2^\omega)$; but this contradicts Q1. □

This corollary admits of a weak converse. Let $\mathrm{AC}_{\mathrm{wo}}$ denote the proposition that every well-ordered set of nonempty sets has a choice function. $\mathrm{AC}_{\mathrm{wo}}$ is very much weaker than AC; for example, although it implies DC, it does not imply $\mathrm{DC}(\aleph_1)$. It is now easy to see that Q1–Q5 plus the additivity property 1.3 can be interpreted in ZF + $\mathrm{AC}_{\mathrm{wo}}$ + $\neg AC$: we say that $Qx(x \in A)$ if A^c is wellorderable. Q1 is

a consequence of $\neg$AC, Q2–Q5 are trivial, and the additivity property follows from AC_{wo}. (We really need some form of choice here because a set can have many wellorderings.) Hence we have proved

2.6. COROLLARY. ZF$\mathcal{Q}$(1–5) + 2.3 *is conservative over* ZF + AC_{wo} + $\neg$AC.

Perhaps this no longer true when we consider full ZF$\mathcal{Q}$.

Corollary 2.4 says that if a set can be "counted" at all, it must be small. Note that it is consistent with ZF$\mathcal{Q}$ that every wellorderable subset of 2^ω is countable; this holds in Solovay's model in which every set of reals is Lebesgue measurable, where we give Q the measure theoretic interpretation. But the consistency proof for ZF$\mathcal{Q}$ shows that the existence of a set of $\aleph_1$ reals (and of a nonmeasurable set) is also consistent with ZF$\mathcal{Q}$. Hence, if CH′ denotes the aleph-free version of the continuum hypothesis, ZF$\mathcal{Q}$ is consistent with both CH′ and $\neg$CH′.

However, as Freiling observed, CH′ is refuted in the system ZF$\mathcal{Q}$ + AC_{wo}! For one can construct a sequence of $\aleph_1$ sets of reals (e.g., by means of the Lebesgue decomposition of $\mathbb{R}$), hence by AC_{wo} there also exists a sequence of reals of length $\aleph_1$. Since in ZF$\mathcal{Q}$ the reals cannot be well-ordered, CH′ must fail.

By keeping track of definability, we obtain

2.7. COROLLARY (ZF$\mathcal{Q}^0$). *Let $\phi(x,\alpha)$ be a formula which may contain additional ordinal and real parameters (but no other parameters). Then $\forall\alpha < \kappa Qx\phi(x,\alpha) \rightarrow Qx\forall\alpha < \kappa\phi(x,\alpha)$. Moreover, if $B =$ Field$(\leq)$ for some definable well-ordering $\leq$, then $Qx(x \notin B)$. For the constructible hierarchy this means the following: if $L(x)$ is the formula "x is constructible," we get $Qx\neg L(x)$.* □

2.8. REMARK. Of course, Q1–Q6 can be eliminated by a set existence axiom of the form "$\exists Q(\varnothing \notin Q \wedge \forall y \in 2^\omega 2^\omega - \{y\} \in Q \wedge \cdots)$", but the Q, the existence of which is asserted, is not unique, so we still need a new primitive. We cannot restate ZF$\mathcal{R}$ in this way, because R is a proper class.

We now investigate the relation between ZF$\mathcal{R}$ and ZF$\mathcal{Q}$. It turns out that ZF$\mathcal{Q}$ and ZF$\mathcal{R}$ are essentially equivalent, and this relationship will allow us to prove the consistency of ZF$\mathcal{Q}$ and the conservativity results mentioned in the previous section. The following results are in part due to Domenico Zambella.

2.9. THEOREM. (i) *ZF$\mathcal{Q}$ can be interpreted in ZF$\mathcal{R}$.*

(ii) *ZF$\mathcal{R}$ can be interpreted in ZF$\mathcal{Q}$.*

PROOF. (i) Let L be the language generated by $\{\in, =\}$. We define a translation * of L(Q) into L(R) as follows:

* is the identity on L formulas

* commutes with $\vee, \wedge, \neg, \rightarrow, \forall, \exists$

$(Qx\phi(x,\vec{y}))^* := \forall x(R(x,\vec{y}) \rightarrow \phi(x,\vec{y})^*)$ (all parameters indicated).

We show that * is a relative interpretation of L(Q) into L(R), i.e., for all ϕ in L(Q): ZF$\mathcal{Q} \vdash \phi$ implies ZF$\mathcal{R} \vdash \phi^*$. (The proof of (ii) will show that * is actually a faithful embedding.)

That * defines a relative interpretation was verified in van Lambalgen [1990], in a different context. We do it again to show the reader which properties of R correspond to quantifier properties. R1 and R5 imply $\forall\vec{y}\exists xR(x,\vec{y})$, hence * is correctly defined as a relative interpretation. The remainder of the argument proceeds by a routine induction on the length of proofs in $\mathcal{Q}$. That the axioms for Q are derivable in $\mathcal{R}$ can be seen as follows:

$(\neg Qx\, x \neq x)^* = \neg\forall x(R(x) \to x \neq x) \leftrightarrow \exists x R(x)$, hence Q1 corresponds to R1 under *.

$(\forall x Qy\, x \neq y)^* = \forall x \forall y(R(y,x) \to x \neq y) \leftrightarrow \forall x \neg R(x,x)$ hence Q2 corresponds to R4 under *.

Q3 holds trivially.

$(Qx\phi(x,\vec{y}) \wedge \forall x(\phi(x,\vec{y}) \to \psi(x,\vec{z})) \to Qx\psi(x,\vec{z}))^* = \forall x(R(x,\vec{y}) \to \phi(x,\vec{y})^*) \wedge \forall x(\phi(x,\vec{y})^* \to \psi(x,\vec{z})^*) \to \forall x(R(x,\vec{z}) \to \psi(x,\vec{z})^*)$; the antecedent implies (using R2) that $\forall x(R(x,\vec{y}\vec{z}) \to \psi(x,\vec{z})^*)$, hence by R5 also $\forall x(R(x,\vec{z}) \to \psi(x,\vec{z})^*)$.

$(Qx\phi(x,\vec{y}) \wedge Qx\psi(x,\vec{z}) \to Qx(\phi(x,\vec{y}) \wedge \psi(x,\vec{z})))^* = \forall x(R(x,\vec{y}) \to \phi(x,\vec{y})^*) \wedge \forall x(R(x,\vec{z}) \to \psi(x,\vec{z})^*) \to \forall x(R(x,\vec{y}\vec{z}) \to \phi(x,\vec{y})^* \wedge \psi(x,\vec{z})^*)$, hence (Q5)* can be derived in $\mathscr{R}$ using R2.

$(QxQy\phi(x,y,\vec{z}) \leftrightarrow QyQx\phi(x,y,\vec{z}))^* = \forall x\forall y(R(x,\vec{z}) \wedge R(y,x\vec{z}) \to \phi(x,y,\vec{z})^*) \leftrightarrow \forall y\forall x(R(y,\vec{z}) \wedge R(x,y\vec{z}) \to \phi(x,y,\vec{z})^*)$, hence (Q6)* can be derived in $\mathscr{R}$ using R6 and R2.

The induction step is almost trivial, since both in the case of $\mathscr{Q}$ and $\mathscr{R}$ the inference rules are those of classical predicate logic. To check the validity of the identity axioms, we have to verify that * commutes with substitution, i.e., that $(Qx\phi(x,y\vec{z}))^*[y = t(\vec{v})] \leftrightarrow (Qx\phi(x,y\vec{z})[y = t(\vec{v})])^*$, but this can be shown using R2, R3 and R5.

(ii) Suppose first that we have a formula $\theta(x,\alpha,\vec{y})$ enumerating (in α) all sets of reals which are definable (in L(Q)!) from $\vec{y}$ and ordinal parameters.

$\hat{}$: L(R) $\to$ L(Q) is defined recursively by

$\hat{}$ is the identity on L,

$R(x,\vec{y})^\wedge = \forall\alpha(Qx\theta(x,\alpha,\vec{y}) \to \theta(x,\alpha,\vec{y}))$,

$\hat{}$ commutes with the logical operators.

We show that $\hat{}$ defines a relative interpretation. Observe that we must have $\forall\vec{y}Qx(R(x,\vec{y})^\wedge)$; for this statement is equivalent to $\forall\vec{y}Qx\forall\alpha(Qx\theta(x,\alpha,\vec{y}) \to \theta(x,\alpha,\vec{y}))$, which follows from the tautology $\forall\vec{y}\forall\alpha(Qx\theta(x,\alpha,\vec{y}) \to Qx\theta(x,\alpha,\vec{y}))$ by Theorem 2.3 and the following property of Q: $Qx(Qx\phi(x) \to \phi(x))$. Hence we have in particular established R1. R2 and R3 are trivial. For R4 observe that $R(x,y)^\wedge$ implies $Qx(x \neq y) \to x \neq y$, which implies $x \neq y$ by Q2.

R5 is a consequence of Q4 and the additivity property 2.3. For suppose $\exists x(R(x,\vec{y}) \wedge \phi(x,\vec{y}))^\wedge$, then by definition $\exists x[\forall\alpha(Qx\theta(x,\alpha,\vec{y}) \to \theta(x,\alpha,\vec{y})) \wedge \phi(x,\vec{y})^\wedge]$. By considering that α for which $\neg\phi(x,\vec{y})^\wedge$ is equivalent to $\theta(x,\alpha,\vec{y})$, we see that we must have $\neg Qx\neg\phi(x,\vec{y})^\wedge$, hence also $\neg Qx\neg(\phi(x,\vec{y})^\wedge \wedge z = z)$. Suppose $\forall x(R(x,z\vec{y}) \to \neg\phi(x,\vec{y}))^\wedge$, then also $\forall x(R(x,z\vec{y})^\wedge \to \neg(\phi(x,\vec{y})^\wedge \wedge z = z))$. If we expand this using the definition of R, we get $\forall x[\forall\alpha(Qx\theta(x,\alpha,z\vec{y}) \to \theta(x,\alpha,z\vec{y})) \to \neg(\phi(x,\vec{y})^\wedge \wedge z = z)]$, whence we get by Q4, $\neg Qx\forall\alpha(Qx\theta(x,\alpha,z\vec{y}) \to \theta(x,\alpha,z\vec{y}))$. However, by definition this is equivalent to $\neg QxR(x,z\vec{y})^\wedge$, and we have seen while verifying R1 that this is false. Lastly, we have to prove the exchange property R6:

We have to show $R(y,\vec{z})^\wedge \wedge R(x,y\vec{z})^\wedge \to R(y,x\vec{z})^\wedge$. Suppose $\neg R(y,x\vec{z})^\wedge$. Then for some α, $\neg(Qy\theta(y,\alpha,x\vec{z})) \to \theta(y,\alpha,x\vec{z}))$. So that by $R(x,y\vec{z})^\wedge$, $\neg Qx\neg(Qy\theta(y,\alpha,x\vec{z}) \to \theta(y,\alpha,x\vec{z}))$, and by $R(y,\vec{z})^\wedge$, $\neg Qy\neg Qx\neg(Qy\theta(y,\alpha,x\vec{z}) \to \theta(y,\alpha,x\vec{z}))$. We apply Q6 and obtain $\neg Qx\neg Qy\neg(Qy\theta(y,\alpha,x\vec{z}) \to \theta(y,\alpha,x\vec{z}))$, whence also

$$\neg Qx\neg(Qy\theta(y,\alpha,x\vec{z}) \wedge Qy\neg\theta(y,\alpha,x\vec{z})).$$

Now $Qy\theta(y,\alpha,x\vec{z}) \wedge Qy\neg\theta(y,\alpha,x\vec{z})$ implies $x \neq x\ y$, Q5 and Q4, so we get $\neg Qx(x = x)$. However, since $\forall x(x = x)$, we also have $Qx(x = x)$, hence there is a contradiction.

It remains to construct the formula $\theta(x,\alpha,\vec{y})$. The obvious way to do this is to copy the construction of the formula defining ordinal definability. To this end, we must extend the reflection theorem to formulas in the language $\mathrm{L}(Q)$, i.e., we have to show:

$$\text{if } \phi \text{ is a formula in } \mathrm{L}(Q), \text{ then } \forall\alpha\exists\beta > \alpha(\phi \leftrightarrow \phi^{V_\beta}).$$

We show how to amend the proof in (Kunen [1980], 137). We first observe that the *quantifier* Q is represented as a *set* in ZF$\mathcal{Q}$; for by the separation axiom we have

$$\exists Q\forall A(A \in Q \leftrightarrow A \subseteq 2^\omega \wedge Qx(x \in A)).$$

We can therefore rewrite a formula $Qx\phi(x)$ as $\exists u \in Q\forall x(x \in u \leftrightarrow \phi(x))$, so we have a reduction to the case of first-order formulas. □

Here is a simple application of 2.9(ii): ZF$\mathcal{R}$ is interpretable in ZF + "all sets of reals are Lebesgue measurable"; for in the latter theory we can interpret ZF$\mathcal{Q}$, and ZF$\mathcal{R}$ is interpretable in ZF$\mathcal{Q}$.

We now restate the above result in terms of derivability.

2.10. COROLLARY. (i) *If* $\phi \in \mathrm{L}(Q)$, ZF$\mathcal{Q} \vdash \phi$ *iff* ZF$\mathcal{R} \vdash \phi^*$,

(ii) *If* $\phi \in \mathrm{L}(R)$, ZF$\mathcal{R} \vdash \phi$ *iff* ZF$\mathcal{Q} \vdash \phi^\wedge$,

(iii) ZF$\mathcal{Q}$ + DC *is consistent.*

PROOF. (i) 2.9(i) shows that if ZF$\mathcal{Q} \vdash \phi$ then also ZF$\mathcal{R} \vdash \phi^*$. To prove the converse, we use the fact, proved in 2.9(i), that ZF$\mathcal{R} \vdash \phi$ implies ZF$\mathcal{Q} \vdash \phi^\wedge$, for ϕ in $\mathrm{L}(R)$. Since R is explicitly definable in ZF$\mathcal{Q}$, we can reproduce a derivation of $(\phi^*)^\wedge$ from ZF$\mathcal{R}$ inside ZF$\mathcal{Q}$, Hence it suffices to show that ZF$\mathcal{Q} \vdash \phi \leftrightarrow (\phi^*)^\wedge$. This is proved by induction. Suppose it holds for ϕ and consider $Qx\phi(x,\vec{y})$. $(Qx\phi(x,\vec{y})^*)^\wedge$ equals $\forall x[\forall\alpha(Qx\theta(x,\alpha,\vec{y}) \to \theta(x,\alpha,\vec{y})) \to (\phi(x,\vec{y})^*)^\wedge]$. By induction, this is equivalent to $\forall x[\forall\alpha(Qx\theta(x,\alpha,\vec{y}) \to \theta(x,\alpha,\vec{y})) \to \phi(x,\vec{y})]$. We have seen above that we get $Qx\forall\alpha(Qx\theta(x,\alpha,\vec{y}) \to \theta(x,\alpha,\vec{y}))$ from 2.3, hence the implication from left to right follows from Q4. To prove the implication in the other direction, assume $Qx\phi(x,\vec{y})$ and $\forall\alpha(Qx\theta(x,\alpha,\vec{y}) \to \theta(x,\alpha,\vec{y}))$; by the induction hypothesis it suffices to show that $\phi(x,\vec{y})$. But this follows by considering that α for which $\phi(x,\vec{y}) \leftrightarrow \theta(x,\alpha,\vec{y})$.

(ii) 2.9(ii) shows that ZF$\mathcal{R} \vdash \phi$ implies ZF$\mathcal{Q} \vdash \phi^\wedge$, and by 2.9(i), ZF$\mathcal{Q} \vdash \phi^\wedge$ implies ZF$\mathcal{R} \vdash (\phi^\wedge)^*$. A trivial induction on length of derivation then shows that ZF$\mathcal{R} \vdash (\phi^\wedge)^*$ implies ZF$\mathcal{R} \vdash \phi$.

(iii) By (i) and 1.6 (b). □

Similarly we have

2.11. THEOREM. (i) ZF$\mathcal{Q}^0$ *can be interpreted in* ZF$\mathcal{R}^0$.

(ii) ZF$\mathcal{R}^0$ *can be interpreted in* ZF$\mathcal{Q}^0$.

PROOF SKETCH. Only the argument for (ii) needs modification. We can no longer conclude immediately from separation that Q is represented as a set in ZF$\mathcal{Q}^0$. We therefore use "bootstrapping": let $\theta_0(x,\alpha,\vec{y})$ denote the formula representing all sets ordinal definable (in $\{\in, =\}$) from $\vec{y}$. Using θ_0, we can define (in ZF$\mathcal{Q}^0$) a set Q^1 by

$$A \in Q^1 \leftrightarrow \exists\alpha(\forall x(x \in A \leftrightarrow \theta_0(x,\alpha,\vec{y})) \wedge Qz\theta_0(z,\alpha,\vec{y})).$$

As in 2.6, we can use Q^1 to show that the reflection theorem holds for L(Q) formulas involving one Q only. This absoluteness can then be used to construct a formula enumerating all sets (of reals), which are definable from ordinals and reals, using formulas that may contain one Q; etc. □

As announced in §1, one can use the argument of Theorem 2.9 to shed light on the role of R6, the Steinitz principle. As we have seen above, R6 corresponds to the Fubini property Q6, which is mainly responsible for the additivity Lemma 2.3.

2.12. THEOREM. (= 1.7.1). *The addition of the axiom* (R7) $R(x,\vec{y}) \rightarrow R(x,\alpha\vec{y})$, *where* α *is an ordinal, is conservative over* ZF$\mathscr{R}$ *for formulas in* $\{\in, =\}$.

PROOF. Evidently, it suffices to construct a translation of ZF$\mathscr{R}$ (1–7) into ZF$\mathscr{R}$ (1–6) which is the identity on formulas in $\{\in, =\}$. First interpret ZF$\mathscr{Q}$ into ZF$\mathscr{R}$ (1–6), as in 2.9(i). Then interpret ZF$\mathscr{R}$ (1–6) into ZF$\mathscr{Q}$; it now suffices to observe that the R' defined in the proof of 2.9(ii) satisfies R7 because of the construction of the formula for ordinal definability. (A similar argument works for ZF$\mathscr{R}^0$.) □

2.13. THEOREM. (= 1.7.3). ZF$\mathscr{R}$ *is conservative* (*in* $\{\in, =\}$) *over* ZF$\mathscr{R}(I)$.

PROOF. ZF$\mathscr{R}(I)$ is the image of ZF$\mathscr{Q}$ under *; now apply 2.10(i). □

It may be of interest to observe that Gödel's principle on ordinals, R7, corresponds to the additivity property 2.3:

2.14. COROLLARY. ZF$\mathscr{Q}$ (1–5) + 2.3 *is interpretable into* ZF$\mathscr{R}$ (1–5, 7) *and conversely*.

PROOF. The proof of 2.9(i) shows that * defines an interpretation of ZF$\mathscr{Q}$ (1–5) + 2.3 into ZF$\mathscr{R}$ (1–5, 7); observe that, once we have R7, the *-translation of $\forall\alpha Qx\phi \rightarrow Qx\forall\alpha\phi$ is trivially derivable! Conversely, that ^ defines an interpretation of ZF$\mathscr{R}$ (1–5, 7) into ZF$\mathscr{Q}$ (1–5) + 2.3 was verified in 2.12. □

2.15. COROLLARY. (= 1.4). *There exists an interpretation of* ZF$\mathscr{R}$ (1–5, 7) *into* ZF + AC_{wo} + "*there is no well-ordering of the reals*".

PROOF. Use Corollaries 2.6 and 2.14. □

§3. Randomness without extensionality? Let us now return to one of the issues touched upon in the introduction, to wit, the question whether randomness really is an extensional notion. If it is, and if one accepts the axioms of §1 or §2, choice fails. But extensionality is used essentially in the proof of Theorem 1.3, since we referred to abstracts like $\{x \mid R(x,g)\}$, which are not available in a nonextensional theory. In this section we look in more detail at the proof of Theorem 1.3, to see where it fails in the absence of extensionality. While this analysis, and related considerations in constructive set theories (where the axiom of choice seems unproblematic if choice functions are required to respect intensional identity only), strongly suggest that ZF$\mathscr{R}$ − EXT + AC is consistent, we have not been able to provide a proof of this conjecture.

First of all, we have to give a description of the intensional set theory we shall consider. (We shall use the term "intensional set theory" to refer to ZF minus extensionality; this should be distinguished from the set theories studied by Myhill and Goodman (for which see Shapiro [1985]). For a good discussion of intensional set theory in our sense, we refer to Beeson [1985].)

In intensional set theory, a set can have many different "names" or "representations" (which can be thought of as definitions). This feature makes it attractive as

a theory in which to formulate randomness, for we now have that randomness is a property of a representation of a sequence and not a property of the sequence *qua* collection of elements. Corresponding to the distinction between a set as a collection of elements and a representation we have two types of identity: intensional identity $\equiv$ (identity between names, or definitional equality) and extensional $=$ (which is of course definable). The presence of two forms of identity has a consequence that several axioms of ZFC can be formulated in various non-equivalent ways.

For instance, if we formulate the pairing axiom using $=$: $\exists x \forall y(y \in x \leftrightarrow y = a \vee y = b)$, we say in effect that we can collect all the names for a and b into one set. We could also consider the weaker from $\exists x(a \in x \wedge b \in x)$: this *is* weaker, because in the absence of extensionality we cannot infer $a \in x \wedge a = a' \rightarrow a' \in x$; in other words, x contains a name for a and b, not necessarily all names. Here, we shall adopt the stronger formulation. Similarly, we have to make a decision about the power set axiom. The usual form says that we can collect all names for subsets of a set b into a new set; we could also consider a weaker form, in which the power set contains a name for each subset of b. Again, we opt for the stronger form, because we would like to say that a random sequence is an element of $\wp(\omega)$.

Choice also presents an interesting example. We can distinguish three versions:

AC($=, \equiv$): the choice function g satisfies $b = b' \rightarrow g(b) \equiv g(b')$,

AC($=, =$): the choice function g satisfies $b = b' \rightarrow g(b) = g(b')$,

AC($\equiv, \equiv$): the choice function g satisfies $b \equiv b' \rightarrow g(b) \equiv g(b')$.

Let ZF$\mathscr{R}^-$ refer to ZF minus extensionality, plus the R axioms, where we allow R in the schemata. Observe that R4 now reads: $R(x, y) \rightarrow x \not\equiv y$.

In this system, we can try to reconstruct the proof of Theorem 1.3 as follows. Let g be a choice function on $\wp(\mathbb{R})$. By separation, $\exists a \forall x(x \in a \leftrightarrow R(x, g))$ and by the definition of a choice function, $\forall a(\forall x(x \in a \leftrightarrow R(x, g)) \rightarrow R(g(a), g)$. Hence, $\exists a \exists y(\forall x(x \in a \leftrightarrow R(x, g)) \wedge g(a) \equiv y \wedge R(y, g))$. To apply R6, we must rewrite this as $\exists y[R(y, g) \wedge \exists a(\forall x(x \in a \leftrightarrow R(x, g)) \wedge g(a) \equiv y)]$ (because $R(y, g)$ does not contain a). R6 yields, for any z, $\exists y[R(y, zg) \wedge \exists a(\forall x(x \in a \leftrightarrow R(x, g)) \wedge g(a) \equiv y)]$. Suppose we pick $z \equiv g(a)$, for some a satisfying $\forall x(x \in a \leftrightarrow R(x, g))$. We then seem to have the following possibilities:

(1) If g satisfies $b = b' \rightarrow g(b) \equiv g(b')$, we have a contradiction as in 1.3;

(2) If g respects extensional $=$ and $R(x, y) \rightarrow x \neq y$, we again have a contradiction as in 2.3;

(3) If there are infinitely many names for $\{x \mid R(x, g)\}$, and if g respects $=$ but $R(x, y) \rightarrow x \not\equiv y$, then there is no contradiction;

(4) Likewise there is no contradiction if there are infinitely many names for $\{x \mid R(x, g)\}$, and if g respects only $\equiv$.

Clearly, the property $b = b' \rightarrow g(b) \equiv g(b')$ for a choice function g is much too strong if the distinction between $=$ and $\equiv$ makes any sense at all. If we review the remaining possibilities, it seems plausible to say that a subset of the reals has infinitely many (different) names; in any case a fixed finite number is implausible. Furthermore, in an intensional approach to randomness R4 should be formulated with $\not\equiv$. Hence, under these assumptions the argument of 1.3 does not go through, irrespective of whether we formulate the axiom of choice with $=$ or with $\equiv$.

We therefore believe that the theory ZF$\mathscr{R}$ minus EXT plus AC is consistent. Of course AC is robbed of much of its strength without extensionality.

§4. Data axioms. So far our axioms have made explicit only very general properties of randomness. For example, the axioms governing Q also allow an interpretation in terms of category: interpret $Qx(x \in A)$ as "A is comeager" (in this case, Q6 corresponds to the Kuratowski-Ulam theorem; see Chapter 16 of Oxtoby [1980]). Most people would think that randomness is connected to measure rather than category, hence some features of randomness are missing. Also, the axioms do not force simple properties like density of random sequences:

4.1. LEMMA. *ZF$\mathscr{Q}$ has models in which the random sequences are not dense.*

PROOF. Let $[0]$ be the cylinder set of all sequences starting with 0. We may suppose $\neg Qx(x \notin [0])$, otherwise we are done. Define a new quantifier $Q^{[0]}$ by $Q^{[0]}\phi \leftrightarrow Qx(x \in [0] \rightarrow \phi)$. Then ZF$\mathscr{Q}$ is also satisfied when we interpret Q as $Q^{[0]}$. Hence in the resulting model there are no random sequences in $[1]$. □

To formulate the additional axioms, which relate randomness to Lebesgue measure, we take our cue from the intuitionistic theory of lawless sequences. A lawless sequence (cf. also the next section) is a process of choosing infinitely many 0's and 1's such that at any stage only finitely many values are known and no restrictions are imposed upon future choices. Two characteristic axioms are:

> Density: every finite binary word is the initial segment of a lawless sequence

and

> Open Data: if we know $A(\alpha)$ for lawless α, this can only be due to our knowledge of an initial segment $\alpha(n)$ of α; hence for all β, if $\beta(n) = \alpha(n)$, then $A(\beta)$.

We will now investigate the analogues of these axioms for random sequences. By now we have four formal systems, and the formulation of the data and density axioms differs for each of them. To spare the reader, we look only at two cases: ZF$\mathscr{Q}$ and ZF$\mathscr{R}^0$.

We first consider ZF$\mathscr{Q}$. As a density axiom we propose

> Density: let A be a Borel set such that $\lambda A > 0$, then $\neg Qx \neg(x \in A)$.

In other words, a set of positive probability contains a random sequence. As a kind of converse, we have

> Inner Data: $Qx(\phi(x) \rightarrow \exists A(A \text{ Borel} \wedge \lambda A > 0 \wedge x \in A \wedge Qz(z \in A \rightarrow \phi(z))))$.

The justification of this principle runs as follows. We think of random sequences as incomplete objects, i.e., objects about which we have only partial information; typically their mode of generation and some initial segment. We now add one more ingredient. We would like our universe of random sequences to be closed under randomness preserving transformations, e.g., subsequence selections. In accordance with usual probability theory, we take a randomness preserving transformation to be a Borel function f: $2^\omega \rightarrow 2^\omega$ such that for all Borel sets B, $\lambda B = 0$ implies $\lambda f^{-1}(B) = 0$.

Now suppose that $\phi(x)$ holds for some random x. Because of the aforementioned closure condition, x is of the form $f(y)$ for measure preserving f and for random y about which we know nothing except, say, an initial segment. But the image of this initial segment under f is the required Borel set of positive measure. We shall see below, in Lemma 4.5, that Inner Data indeed allows us to prove closure under randomness preserving transformations. (Actually, Inner Data resembles the intuitionistic principle "Analytic Data" (Troelstra [1977], 78) more than Open Data. In any case it is clear that, by resorting to a specimen of intuitionistic reasoning in the above justification, we have given a more or less constructive meaning to the dual quantifier $\neg Q \neg$.)

We next consider ZF$\mathscr{R}^0$. We now have to take parameters into consideration. Suppose $\phi(x) = \phi(x, \vec{y})$, where all parameters are real (by R7, we may forget about additional ordinal parameters). We then set

Inner Data^0: $\forall x(R(x,\vec{y}) \wedge \phi(x,\vec{y}) \rightarrow \exists A(\lambda A > 0 \wedge x \in A \wedge \forall z(R(z,\vec{y}) \wedge z \in A \rightarrow \phi(z,\vec{y}))))$.

Density becomes

Density^0: For Borel sets A, $\lambda A > 0 \rightarrow \forall \vec{y} \exists x(x \in A \wedge R(x, \vec{y}))$.

As we shall see in the proof of 4.2, these statements express truths about Solovay forcing. The reader may also wish to compare the above formulation of Inner Data with the parametrized form of the axiom of Open Data, given in §5.

Henceforth we abbreviate Inner Data to ID, and Density to D; if we consider these axioms in the context of ZF$\mathscr{R}^0$ we add the superscripts. When we add both D$^{(0)}$ and ID$^{(0)}$ to a theory we indicate this by IDD$^{(0)}$.

4.2. Lemma. ZF$\mathscr{R}^0$ + AC + IDD^0 *holds in the model obtained by generically adding* $\kappa \geq \omega_1$ *random reals to a groundmodel satisfying* V = L.

Proof. Let $\mathscr{M}[\mathscr{G}]$ be the generic extension constructed in Lemma 1.5. It suffices to show that Density and Inner Data hold in $\mathscr{M}[\mathscr{G}]$. Density follows from the properties of Solovay randomness: if for some $\vec{y}$, $\forall x(R(x,\vec{y}) \rightarrow x \notin A))$, then since $\{x \mid R(x,\vec{y})\}$ has outer measure 1, $\lambda A = 0$. Moreover, ID was already implicitly verified in the proof of R5. For we showed that, given a formula $\phi(x,\vec{y})$, there exists a Borel set A with code in $\mathscr{M}[\vec{y}]$, such that for all x random over $\mathscr{M}[\vec{y}]$, $x \in A$ iff $\mathscr{M}[\mathscr{G}] \models \phi(x,\vec{y})$. If $\phi(x,\vec{y})$ is verified for some random x, $\lambda A > 0$. Hence we have shown

$$\forall x(R(x,\vec{y}) \wedge \phi(x,\vec{y}) \rightarrow \exists A(\lambda A > 0 \wedge x \in A \wedge \forall z(R(z,\vec{y}) \wedge z \in A \rightarrow \phi(z,\vec{y})))). \qquad \square$$

As in §2 we obtain

4.3. Corollary. ZF$\mathscr{Q}$ + DC + IDD *is consistent.*

A consequence of ID that is useful in applications is given in

4.4. Lemma (ZF$\mathscr{Q}$ + DC + ID). *For any formula* $\phi(x)$ *there is a Borel set* A *such that* $Qx(x \in A \leftrightarrow \phi(x))$.

Proof. We may assume $\neg Qx \neg \phi(x)$, for otherwise we can take $A = \varnothing$. Consider the set $\{A \mid \lambda A > 0 \wedge Qx(x \in A \rightarrow \phi(x))\}$, which exists in ZF$\mathscr{Q}$. Let A_0 be an element of maximal measure. A_0 exists because $s := \sup\{\lambda A \mid \lambda A > 0 \wedge Qx(x \in A$

$\rightarrow \phi(x))\} \leq 1$; choose (DC!) $A_1, \ldots, A_n, \ldots$ such that $\lambda A_n \rightarrow s$, and put $A_0 := \bigcup A_n$. Then we must have $Qx(x \in A_0 \leftrightarrow \phi(x))$, for otherwise $\neg Qx \neg(\phi(x) \wedge x \notin A_0)$ and ID shows that A_0 was not maximal after all. □

The reader will have noticed that the proofs of R5 and ID^0 in the random real extension are very similar. It may therefore be instructive to observe that ID^0 can be derived from two weaker axioms, with the help of R5. The axioms are

(1) Let B be a Borel set with Borel code u, such that $\lambda B = 1$. Then $\forall x(R(x,u) \rightarrow x \in B)$.

(2) For any formula $\phi(x, \vec{y})$, there exists a Borel set B such that $\forall x(R(x, \vec{y}) \rightarrow (\phi(x, \vec{y}) \leftrightarrow x \in B))$.

Suppose $R(x, \vec{y}) \wedge \phi(x, \vec{y})$. To prove ID^0, it suffices to show that for the B given by (2), $\lambda B > 0$. Suppose $\lambda B = 0$. Let B have Borel code u. By (1), $\forall x(R(x,u) \rightarrow x \notin B)$, whence by R2, $\forall x(R(x, u\vec{y}) \rightarrow x \notin B)$. By (2), $\forall x(R(x, u\vec{y}) \rightarrow \neg \phi(x, \vec{y}))$. We may now apply (the contraposition of) R5 to obtain $\forall x(R(x, \vec{y}) \rightarrow \neg \phi(x, \vec{y}))$, a contradiction.

Conversely, in $\text{ZF}\mathscr{R}^0$ (minus R5) + D^0, ID^0 implies (1), (2) and R5. (2) follows from 4.4. To prove (1), let B be a Borel set with code u such that $\lambda B = 1$ and suppose that for some x, $R(x, \vec{y}) \wedge x \notin B$. By ID^0, there exists A such that $\lambda A > 0$ and $\forall z(R(z, \vec{y}) \wedge z \in A \rightarrow z \notin B)$. We can rewrite this as $\forall z(R(z, \vec{y}) \rightarrow (z \notin B \vee z \notin A))$; but since $\lambda(B^c \cup A^c) < 1$, we have a contradiction with D^0.

Lastly, we prove R5. Suppose $\exists x(R(x, \vec{y}) \wedge \phi(x, \vec{y}))$ and $\forall x(R(x, z\vec{y}) \rightarrow \neg \phi(x, \vec{y}))$. ID^0 gives us a Borel set A such that $\lambda A > 0$ and $\forall x(R(x, \vec{y}) \wedge x \in A \rightarrow \phi(x, \vec{y}))$. By hypothesis, $\forall x(R(x, \vec{y}) \wedge x \in A \rightarrow \neg R(x, z\vec{y}))$, which is equivalent to $\forall x(R(x, \vec{y}) \wedge R(x, z\vec{y}) \rightarrow x \notin A)$. By R2 we obtain $\forall x(R(x, z\vec{y}) \rightarrow x \notin A)$, which conflicts with D^0.

The presence of (1) of course explains why we had to use Solovay forcing.

The next lemma states a weak type of closure necessary to prove measure extension theorems:

4.5. Lemma ($\text{ZF}\mathscr{Q}$ + IDD). *Let f: $2^\omega \rightarrow 2^\omega$ be a bijective measure preserving Borel function. Then $Qx\phi(x) \leftrightarrow Qx\phi(f(x))$.*

Proof. Suppose $Qx\phi(x) \wedge \neg Qx\phi(f(x))$. By Inner Data, we obtain a Borel set A such that $\lambda A > 0$ and $Qx(x \in A \leftrightarrow \neg \phi(f(x)))$. Then $f[A]$, the image of A under f, is Borel and we have $Qx(x \in f[A] \leftrightarrow \neg \phi(x))$; but since $Qx\phi(x)$, we get $Qx(x \notin f[A])$, whence $\lambda f[A] = 0$; a contradiction. For the other direction, consider f^{-1} (which is also Borel and measure preserving). □

We are now ready for the main result of this section.

4.6. Theorem ($\text{ZF}\mathscr{Q}$ + DC + IDD). *There exists a translation invariant extension μ of Lebesgue measure to all of $\wp(2^\omega)$. This μ can be computed as follows: every $B \in \wp(2^\omega)$ can be written as $A \Delta N$, where A is Borel and N is a μ-nullset, and $\mu A \Delta N = \lambda A$.*

Proof. The idea that Fubini's theorem is connected to extensions of Lebesgue measure is suggested by Simms [1989]; the consistency result itself was already obtained by Solovay. Let $\mathscr{A}$ denote the Borel σ-algebra on 2^ω. Let $\mathscr{N}$ denote $\{B \in \wp(2^\omega) \mid Qx(x \notin B)\}$. By Theorem 1.3, $\mathscr{N}$ is closed under countable unions. By Q4, $\mathscr{N}$ is closed under subsets. Hence $\mathscr{N}$ is a σ-ideal. Let $\mathscr{B} = \{A \Delta N \mid A \in \mathscr{A}, N \in \mathscr{N}\}$. It is easy to see that $\mathscr{B}$ is a σ-algebra. We would like to define μ on $\mathscr{B}$ such that $\mu A \Delta N = \lambda A$. To make μ well defined, we should have $A_1 \Delta N_1 = A_2 \Delta N_2$ implies $\lambda A_1 = \lambda A_2$. But $A_1 \Delta N_1 = A_2 \Delta N_2$ implies $A_1 \Delta A_2 = N_1 \Delta N_2$ and since $N_1 \Delta N_2 \in$

$\mathcal{N}$, we get $Qx(x \notin A_1 \Delta A_2)$. By Inner Density, $\lambda A_1 \Delta A_2 = 0$. Hence μ is well-defined and is easily shown to be σ-additive. Moreover, $N \in \mathcal{N}$ implies $\mu N = 0$.

We now show that $\mathscr{B} = \wp(2^\omega)$. Choose $B \in \wp(2^\omega)$. Apply Lemma 4.4 to get a Borel A such that $Qx(x \in A \leftrightarrow x \in B)$, i.e., $Qx(x \notin A \Delta B)$. Hence $A \Delta B \in \mathcal{N}$. We may now write $B = A \Delta (A \Delta B)$, which shows that $B \in \mathscr{B}$.

That μ is translation invariant follows from Lemma 4.5. □

As is well-known, Theorem 4.6 implies the nonexistence of ultrafilters on $\wp(\omega)$, the nonexistence of Hamelbases, the existence of uncountable cardinals which carry σ-complete ultrafilters, and the failure of the axiom of choice for uncountable sets of two element sets. It is perhaps instructive to compare Theorem 4.6 to the famous measure extension theorem of Oxtoby and Kakutani [1950]. They showed that, in the presence of AC, Lebesgue measure can be extended to a translation invariant measure on a σ-algebra whose smallest set of generators has cardinality $2^{2^{\aleph_0}}$, so that in a sense the Lebesgue algebra is very small. Theorem 4.6 points in the opposite direction: in ZF$\mathscr{Q}$ + DC + IDD, the Lebesgue algebra is a very large subalgebra of $\wp(2^\omega)$. (It is doubtful, however, whether these results are relevant for the immediate concerns of, say, probabilists; in practice, one needs only the fact that the σ-algebras under investigation are closed under a restricted number of operations.)

4.7. Lemma (ZF$\mathscr{Q}$ + DC + IDD). *μ is κ-additive, for any ordinal κ.*

Proof. Suppose $\{A_\alpha\}_{\alpha<\kappa}$ are pairwise disjoint subsets of 2^ω. At most countably many of them can have positive μ measure; say these are $\{A_n\}_{n<\omega}$. Then

$$\sum_{\alpha<\kappa} \mu A_\alpha = \sum_{n<\omega} \mu A_n + \mu \cup \{A_\alpha \mid \omega \leq \alpha < \kappa\};$$

by the additivity property 1.3, the last term equals 0. □

The meaning of these results can perhaps best be explained by referring to two rivalling traditions in probability theory. In the approach universally used today (associated with the name of Kolmogorov), a probability is taken to be a σ-additive measure on a σ-algebra. Other notions that are of interest to a probabilist, foremost among them independence, are defined in terms of measure. But there is an older tradition, now nearly extinct, in which rather randomness and independence are taken as primitive. This approach goes back to von Mises, who tried to capture the independence inherent in random sequences by means of axioms about subsequence selection. Although these axioms themselves are fairly unwieldy, it still seems to us that the general idea of treating fundamental probabilistic notions as primitives has some potential. For instance, we have seen just now that the whole business about σ-algebras actually becomes superfluous. This may not be immediately apparent, because we formulated D and ID in terms of Borel sets and Lebesgue measure; but a glance at the axioms will show that we could have used closed sets instead, and *their* Lebesgue measure is easily computed.

4.8. Digression: stronger forms of inner data. We could further effectivize the quantifier Q by building some uniformity into ID, for instance as follows:

$$\text{ID}^2\colon QxQy(\phi(x,y) \rightarrow \exists A(A \text{ Borel} \wedge \lambda^2 A > 0 \wedge \langle x,y\rangle \in A \wedge QzQz'(\langle z,z'\rangle \in A \rightarrow \phi(z,z'))).$$

ID^2 also holds in the model considered above. We can formulate ID^2 equivalently (and perhaps more perspicuously) as a uniform version of ID: for any formula ϕ there is a set-valued Borel measurable mapping F such that $(*)$ $Qx(\phi(x,y) \to \lambda F(y) > 0 \wedge x \in F(y) \wedge Qz(z \in F(y) \to \phi(z,y)))$. ID^2 implies the existence of such an F by setting $F(y) := A_y$; conversely, we obtain A by applying ID to $(*)$. By considering higher dimensional versions of ID, we can replace y by $\vec{y}$. Because of the uniformity present in ID^2, this axiom seems to be stronger than ID, but we have been unable to obtain a proof.

ID^2 rules out constructions such as Davies's theorem [1963], which, assuming AC, establishes the existence of a real function f, such that the plane is the union of countably many congruent copies of f. For by Q2 and Q4, $QxQy(y \notin f(x))$, and ID^2 yields the existence of an isometry invariant σ-additive extension of λ^2 to all subsets of the plane in the same manner as above.

§5. Excursion into intuitionism: lawless sequences. In the previous sections we have repeatedly compared random sequences with lawless sequences in order to justify axioms for the former. In this section we show that there is in fact a precise correspondence between the two: we shall define an interpretation of $\mathrm{ZF}\mathscr{Q}^0$ into intuitionistic set theory enriched with axioms for lawless sequences. We shall need some groundwork to introduce the relevant systems, but the result itself is entirely trivial.

We begin with a brief exposition of the axioms for lawless sequences; the interested reader can find a more detailed treatment in Troelstra [1977] or in Troelstra and van Dalen [1988].

Lawless sequences are processes of assigning (in our case) 0 or 1 to the arguments $1, 2, 3, \ldots$, such that (1) at any stage only finitely many values are known and (2) at no stage the possibility for choosing values is restricted.

We expand intuitionistic set theory with variables for lawless sequences, denoted as $\alpha, \beta, \ldots$. We use w as a variable over finite binary sequences. $[w]$ is the set of infinite binary sequences which have initial segment w.

$\#(\alpha, \beta_1, \ldots, \beta_n)$ is defined as $\alpha \neq \beta_1 \wedge \cdots \wedge \alpha \neq \beta_n$; and the quantifier $\underline{\forall}\gamma B(\gamma, \beta_1, \ldots, \beta_n)$ is defined as $\forall\gamma(\#(\gamma, \beta_1, \ldots, \beta_n) \to B(\gamma, \beta_1, \ldots, \beta_n))$.

The axioms are as follows:

LS1. $\forall w \exists \alpha(\alpha \in [w])$,

LS2. $\alpha = \beta \vee \alpha \neq \beta$,

LS3. $A(\alpha, \beta_1, \ldots, \beta_n) \wedge \#(\alpha, \beta_1, \ldots, \beta_n) \to \exists w(\alpha \in [w] \wedge \underline{\forall}\gamma \in [w] A(\gamma, \beta_1, \ldots, \beta_n)$.

We add a few words of explanation. LS1 says that we are allowed to specify an initial segment in advance. LS2 is based on the fact that extensional and intensional identity coincide for lawless sequences. Let $\equiv$ denote intensional identity: $\alpha \equiv \beta$ means that α and β are given to us as the same process. Then obviously $\alpha \equiv \beta \vee \neg\alpha \equiv \beta$. It is, however, easy to show that for extensional identity $=$, $\alpha = \beta \leftrightarrow \alpha \equiv \beta$; the implication from right to left is trivial and the implication from left to right holds in virtue of the intuitionistic meaning of the implication: given the nature of lawless sequences, we can only have a proof of $\alpha = \beta$ if we already know that $\alpha \equiv \beta$. Inequality can therefore be seen as a kind of independence: $\alpha \neq \beta$ holds if

and only if α and β are given to us as different processes; but since choosing lawlessly means that there are no restrictions, the process generating β in fact does not influence the process generating α. It will therefore come as no surprise that in the interpretation constructed below, independence is interpreted as inequality.

LS3 is clearly inherent in the meaning of lawlessness, especially if we consider the parameterfree version: $A(\alpha) \to \exists w(\alpha \in [w] \wedge \forall \gamma \in [w] A(\gamma))$. For if we have a proof of $A(\alpha)$, this can only be on the basis of a finite segment of α. In the general case of formulas with parameters, we have to add some provisos (like the condition $\#(\alpha, \beta_1, \ldots, \beta_n)$) to avoid inconsistency. For instance, if we formulated Open Data as $A(\alpha, \beta_1, \ldots, \beta_n) \to \exists w(\alpha \in [w] \wedge \forall \gamma \in [w] A(\gamma, \beta_1, \ldots, \beta_n)$, we would obtain an immediate contradiction for the formula $\alpha = \beta$ (cf. also the formulation of the Inner Data axiom in §4.) In the formula $A(\alpha, \beta_1, \ldots, \beta_n)$ we have exhibited only the parameters $\beta_1, \ldots, \beta_n$ which are assumed to be lawless; $A(\alpha, \beta_1, \ldots, \beta_n)$ *may contain parameters for, and quantifiers over arbitrary sets.* In this context, Troelstra remarks: "LS3 seems also to be justified for predicates containing ... parameters for nonlawlike objects not constructed from lawless sequences ..., provided extensionality holds w.r.t. all function and set parameters." (Troelstra [1977], p. 29)

We now turn to intuitionistic set theory. IZF (cf. Friedman [1973]) is the following set of axioms, formulated in intuitionistic logic:

1. (Extensionality) $x = y \wedge x \in w \to y \in w$,
2. ($\in$-induction) $\forall x(\forall u \in x\phi(u) \to \phi(x)) \to \forall x\phi(x)$,
3. (Separation) $\exists y\forall z(z \in y \leftrightarrow z \in x \wedge \phi)$,
4. (Pairs) $\exists y(u \in y \wedge v \in y)$,
5. (Union) $\exists y\forall z(\exists u \in x(z \in u) \to z \in y)$,
6. (Collection) $\forall x \in u\exists y\phi \to \exists v\forall x \in u\exists y \in v\phi$,
7. (Power set) $\exists y\forall z(\forall u(u \in z \to u \in x) \to z \in y)$,
8. (Infinity) $\exists y(\exists z(z \in y) \wedge \forall x \in y\exists z \in y(x \in z))$.

$\neg\phi$ is defined as $\phi \to \bot$, where $\bot$ is some absurdity; e.g., $\bot = \forall xy(x \in y)$. Furthermore, equality is *defined* as $x = y \Leftrightarrow \forall z(z \in x \leftrightarrow z \in y)$.

Let IZF$\mathscr{L}\mathscr{S}$ be the above set of axioms, together with LS1–3. One comment is in order: it might be thought that the presence of lawless sequences destroys extensionality, but that is not so: as was pointed out while justifying LS2, we only have a proof of $\alpha = x$ when α and x are given to us as the same object.

5.1. Theorem. *ZF$\mathscr{Q}$ can be interpreted in* IZF$\mathscr{L}\mathscr{S}$.

Proof. The interpretation is a slight variant of Friedman's negative translation [1973]. Friedman proceeds as follows: he first embeds ZF into ZF$^-$ (that is, ZF minus the axiom of extensionality) and then extends Gödel's negative translation to embed ZF$^-$ into IZF$^-$.

The first part can be taken over unchanged. For the second part, we have to extend the negative translation with a clause for Q. This is most easily achieved when we formulate ZF$\mathscr{Q}$ in such a way that there is a typographical distinction between variables bound by $\forall$ and variables bound by Q. So suppose the variables bound by Q are represented by lower case Greek letters. Then we have:

$(x \in y)^* = \neg\neg x \in y$,

$(\neg\phi)^* = \neg\phi^* := \phi^* \to \bot$,

$(\phi \vee \psi)^* = \neg(\neg\phi^* \wedge \neg\psi^*)$,

$(\phi \wedge \psi)^* = \phi^* \wedge \psi^*$,
$(\phi \rightarrow \psi)^* = \phi^* \rightarrow \psi^*$,
$(\forall x\phi)^* = \forall x\phi^*$,
$(\exists x\phi)^* = \neg\neg\exists x\phi^*$,
$(Q\gamma\phi(\gamma, \vec{y}))^* = \underline{\forall}\gamma\phi(\gamma, \vec{y})^*$ *where* $\underline{\forall}$ *refers only to the Greek letters in* ϕ.

The Q-clause for the translation is justified by Troelstra's remark cited above. All Q axioms except Q4 are trivially true under this interpretation. To verify Q4, we shall use the axioms of density and open data (the reader will note the analogy with the proof of R5 from ID^0 and D^0). We only exhibit lawless parameters. So suppose we have $\forall\alpha(\alpha \neq \beta, \gamma \rightarrow \phi^*(\alpha, \gamma))$ (β not free in ϕ^*), then we have to show $\forall\alpha(\alpha \neq \gamma \rightarrow \phi^*(\alpha, \gamma))$ (cf. the proof in 2.9 that Q4 is interpreted by R5).

If $\beta = \gamma$ we are done. So (LS1) suppose $\alpha \neq \gamma$, $\beta \neq \gamma$. By applying LS3 to β in $\forall\alpha(\alpha \neq \beta, \gamma \rightarrow \phi^*(\alpha, \gamma))$ we get

$$\exists w(\beta \in [w] \wedge \forall\beta' \in [w](\beta' \neq \gamma \rightarrow \forall\alpha(\alpha \neq \beta', \gamma \rightarrow \phi^*(\alpha, \gamma)))).$$

By LS2 we can take $\beta' \in [w]$ such that $\alpha \neq \beta'$ and $\gamma \neq \beta'$ and $\phi^*(\alpha, \gamma)$. This completes the proof of the interpretation of Q4. □

5.2. Remark. By intuitionistic logic, we also have $\neg\neg\exists\alpha(\alpha \neq \beta \wedge \phi(\alpha, \beta)) \rightarrow \neg\neg\exists\alpha\,(\alpha \neq \beta, \gamma \wedge \phi(\alpha, \beta))$ (γ not free in ϕ), so that the $\neg\neg$-translation of R5 follows from Open Data. This is the motivation behind the remark in the second paragraph of §1 that, whereas the ordinary reflection principle reflects properties of the classical universe V downward, R5 reflects properties on the intuitionistic universe. Given this close connection between $\mathrm{ZF}\mathscr{Q}$ and $\mathrm{IZF}\mathscr{L}\mathscr{S}$, is there a way to see the failure of choice in both cases as an instance of the same phenomenon?

§6. Appendix: related approaches. Here we discuss two other attempts to add axioms for randomness to set theory, those of Myhill and Freiling.

6.1. Myhill's axioms. These were alluded to in the quotation from Kreisel in the introduction. They are reported in paragraph 10 of Kruse [1967].

Again, let 2^ω be the space of infinite binary sequences and let λ denote Lebesgue measure on this space, i.e., $(\frac{1}{2}, \frac{1}{2})^\omega$. Let $R(x)$ be a predicate which should intuitively be interpreted as "x is random." Myhill tries to formalize the intuition that random sequences should satisfy "all" properties of probability 1 and no "special" properties.

His axioms (M) are

(M1) $\lambda\{x \in 2^\omega \mid R(x)\} = 1$;

(M2) If $\phi(x)$ is a formula in one free variable, then $\lambda\{x \in 2^\omega \mid \phi(x)\} = 1$ implies $\forall x(R(x) \rightarrow \phi(x))$.

The restriction that ϕ contain no parameter beside x is obviously necessary, since otherwise we could take the formula $x \neq y$. The formulation of (M2) contains a deliberate ambiguity, however: is R allowed to occur in ϕ or not? If not, then M is obviously conservative over ZFC. On the other hand, if we do allow R in ϕ, then R need no longer be explicitly definable. Myhill comments:

> [This] would accord with a prejudice of mine which I derived from Feller, i.e. that randomness is an intensional notion, not definable in the usual

> mathematical terms. The "circularity" of the schema above with R allowed to appear in ϕ is quite justified if we are convinced that R belongs to a new order of ideas, entirely outside the set theoretic order.

Myhill's emphasis is somewhat different from ours: it will be clear by now that we believe that randomness is a natural continuation of set theoretic ideas (cf. the quotation from Bernays in the introduction) instead of being "entirely outside the set theoretic order."

An immediate consequence is that M is no longer conservative over ZFC, for it implies the by-now-familiar consequence that there is no definable well-ordering of the continuum: if there were such a well-ordering, then we could define the least random element, in contradiction with (M2). It is clear that this proof works only when R is also allowed to occur in the schemata of ZF. Note that, on the liberal interpretation of (M2), the consistency of ZFC + M (with R allowed in the schemata of ZF) is not immediate.

6.1.1. LEMMA. *ZFC + M is consistent.*

PROOF SKETCH. We use a model of Stern [1985], Theorem 2(v). Starting from a model $\mathscr{M}$ of ZF + V = L, Stern constructs a generic extension in which all ordinal definable sets are Lebesgue measurable, and in which the set of Solovay random reals (over $\mathscr{M}$) has measure 1. Hence if we interpret $R(x)$ as: x is Solovay random over $\mathscr{M}$, (M1) holds. Furthermore, R is definable, hence to verify (M2) it suffices to show that a definable set of reals with measure 1 contains all random reals. This is established in Stern [1985], §4.7. □

In view of the theory of randomness outlined in §1, the following modification of Myhill's axioms seems reasonable: we have a relation $R(x, \vec{y})$ (with real parameters) such that

(M0) $R(x, \vec{y})$ satisfies $\mathscr{R}^0$,

(M1) $\forall \vec{y} \lambda\{x \in 2^\omega \mid R(x, \vec{y})\} = 1$,

(M2) If $\lambda\{x \in 2^\omega \mid \phi(x, \vec{y})\} = 1$, then $\forall x(R(x, \vec{y}) \rightarrow \phi(x, \vec{y}))$.

These axioms can be interpreted in Solovay's model for "all sets of reals are Lebesgue measurable." Is it possible to prove the consistency of this theory without using an inaccessible?

6.2. Freiling and the continuum hypothesis. We will next investigate the relation of Freiling's axioms of symmetry [1986] to the axioms introduced here. Freiling writes:

> Suppose we were to throw a random dart at the real number line and ask whether the dart landed on a rational number. The outcome is, of course, predictable. We could say in advance that the dart will (with probability one) land on an irrational number. Furthermore, let us agree that the reason does not depend on any particular property of the set of rational numbers except that it is countable and its members are determined before we make our throw.
>
> Now suppose we were to throw two darts and ask whether the second dart was a rational multiple of the first one. The answer would likewise be no, since by the time we throw the second dart there are only countably many points which it has to miss, and membership in this countable set is predetermined by the first dart.

Suppose then that we have a function $f\colon \mathbb{R} \to \mathbb{R}_{\aleph_0}$ (i.e., f assigns to each real a countable set of reals). The second dart will not be in the countable set assigned to the first dart. Now by the symmetry of the situation (the real line does not know which dart was thrown first or second), we could also say that the first dart will not be in the set assigned to the second. This leads us to the following natural proposition:

$$A_{\aleph_0} \qquad \forall f\colon \mathbb{R} \to \mathbb{R}_{\aleph_0} \exists xy(x \notin f(y) \wedge y \notin f(x)),$$

the intuition being that x and y could be found by independently throwing two random darts.

He then proceeds to prove that $A_{\aleph_0}$ is equivalent to $\neg$CH. We intend to show here that Freiling's intuitive motivation for $A_{\aleph_0}$ is entirely captured by $\mathcal{Q}$, in the following sense: a suitable fragment of ZF$\mathcal{Q}$ (basically one allows two iterations of Q only) suffices to derive $A_{\aleph_0}$ and, conversely, that fragment can be interpreted in ZFC + $A_{\aleph_0}$.

Although the resulting fragment is admittedly ad hoc, we shall try to motivate it by referring back to the proof of Theorem 1.3. It will be observed that, in that proof, we needed only statements of the form $Qx(\langle x, y\rangle \in U)$ or $QxQy(\langle x, y\rangle \in \mathrm{V})$. Accordingly, we can define a class of *elementary statements* as follows.

6.2.1. Definition. The class of n-elementary statements is defined by

(1) $\langle x_1, \ldots, x_n\rangle \in U$, where $U \subseteq (2^\omega)^n$, is n-elementary,

(2) n-elementary statements are closed under $\to, \neg, \wedge, \vee, \forall, \exists$,

(3) if ϕ is elementary and not of the form $\forall x\psi$ or $\exists x\psi$, then $Qx\phi$ is n-elementary.

ZF$\mathcal{E}_n$ is obtained from ZF$\mathcal{Q}$ by applying the following restrictions:

(1) the Q axioms are formulated for n-elementary statements only,

(2) we allow only n iterations of Q,

(3) in n-elementary statements, we allow only those $U \subseteq (2^\omega)^n$ for which $\{x \mid \langle x_1, \ldots, x_{n-1}, x\rangle \in U\}$ is countable for all $\langle x_1, \ldots, x_{n-1}\rangle$.

We apologize for the lack of elegance, both of the preceding definitions and of the proof of the main result, 6.2.7. We include this material only because it shows that Freiling's intuitions fit squarely in the framework of the preceding sections. Furthermore, it may be of interest to see that Q can also have a cardinality interpretation and that the size of the continuum is related to the number of iterations of Q.

6.2.2. Theorem. (ZF$\mathcal{E}_2$) $A_{\aleph_0}$.

Proof. Suppose we are given $f\colon 2^\omega \to (2^\omega)_{\aleph_0}$. It suffices to show $QxQy(x \notin f(y) \wedge y \notin f(x))$, for generally $Qx\phi$ implies $\exists x\phi$. By Q6 and Q5, we only have to prove $QxQy(y \notin f(x))$. However, this follows from $\forall xQy(y \notin f(x))$, which is a consequence of 1.3. □

Clearly, Freiling's symmetry principle ("the real line does not know the order of the darts") corresponds to Q6.

6.2.3. Corollary. (ZF$\mathcal{E}_2$) $\neg$CH.

Proof. If CH, there is a well-ordering $\leq$ of 2^ω of length $\aleph_1$. Now consider $f(y) := \{x \mid x \leq y\}$. Then $\forall x\forall y(x \in f(y) \vee y \in f(x))$; but since each $f(y)$ is countable we get a contradiction from $A_{\aleph_0}$. □

Actually something stronger holds. Let $(2^\omega)_n$ denote the set of n-element subsets of x. Since $(2^\omega)_n$ can be coded into 2^ω, we can also consider functions $f\colon (2^\omega)_n \to (2^\omega)_{\aleph_0}$. This means that we are able to formulate and prove a generalization of $A_{\aleph_0}$ in our set up. First a

6.2.4. DEFINITION. Let $f\colon (2^\omega)^n \to (2^\omega)_{\aleph_0}$. A set $X \subseteq 2^\omega$ is called f-incomparable if for any n distinct elements $x_1, \dots, x_n$ of X, $x_n \notin f(\{x_1, \dots, x_{n-1}\})$.

The generalization of $A_{\aleph_0}$ can then be formulated as follows:

$A^n_{\aleph_0}$ $\forall f\colon (2^\omega)_{n-1} \to (2^\omega)_{\aleph_0}$ $\exists f$-incomparable set of size n.

6.2.5. THEOREM (ZF$\mathscr{E}_n$). *For all* n, $A^n_{\aleph_0}$.

PROOF. Along the same lines as for $n = 2$. To show how Q6 is applied we do the case $n = 3$. Choose f. By 1.3 $\forall x \forall y Qz(z \notin f(\{x,y\})$, hence by monotonicity $QxQyQz(z \notin f(\{x,y\})$. By Q6, we get $QzQyQx(z \notin f(\{x,y\})$, whence $QxQyQz(x \notin f(\{z,y\})$ by Q3; and similarly $QxQyQz(y \notin f(\{x,z\})$. Now apply Q5. □

Freiling proves

6.2.6. THEOREM (ZFC). $A^n_{\aleph_0}$ *is equivalent to* $2^{\aleph_0} \geq \aleph_n$.

We will now establish an equivalence between Freiling's approach and ours.

6.2.7. THEOREM. ZF$\mathscr{E}_2$ + AC *is interpretable in* ZFC + $A_{\aleph_0}$.

PROOF. The obvious interpretation of Q would be to put A in Q if A is co-countable, but this runs afoul of Q6. We therefore have to give a contextual definition of Q, depending on an additional parameter. Indeed, it seems impossible to give a *uniform* interpretation of Q in the absence of strong hypotheses like, e.g., "all sets of reals have the Baire property."

Let $\mathscr{A}$ be the σ-algebra of countable (i.e., finite or countably infinite) and co-countable sets. Define $\nu\colon \mathscr{A} \to \{0,1\}$ by $\nu A = 1$ iff A is co-countable. ν is σ-additive. We show how to extend the product measure $\nu \times \nu$ on $\mathscr{A} \times \mathscr{A}$ to a measure μ on an extension of $\mathscr{A} \times \mathscr{A}$ that will serve to interpret Q. Expressions like "full", "null" or "almost all" refer to ν.

We first need a lemma which shows that Sierpinski's counterexample to a Fubini theorem without joint the measurability condition does not exist under $A_{\aleph_0}$.

6.2.8. LEMMA (ZFC + $A_{\aleph_0}$). *There is no set that is null for almost all horizontal sections and full for almost all vertical sections.*

PROOF. See Freiling [1986], p. 197. □

If $A \subseteq 2^\omega \times 2^\omega$, define $A_x := \{y \mid \langle x,y\rangle \in A\}$ (vertical section) and $A^y := \{x \mid \langle x,y\rangle \in A\}$ (horizontal section). Let $\mathscr{N}_x$ denote $\{A \subseteq 2^\omega \times 2^\omega \mid \forall x \in 2^\omega\, A_x$ is null$\}$ and similarly $\mathscr{N}^y := \{A \subseteq 2^\omega \times 2^\omega \mid \forall y \in 2^\omega\, A^y$ is null$\}$. Obviously $\mathscr{N}_x$ and $\mathscr{N}^y$ are σ-ideals.

6.2.9. LEMMA (ZFC + $A_{\aleph_0}$). *Let A be $\nu \times \nu$ measurable and suppose $A \subseteq \bigcup B_i$, where $\{B_i\}$ is a countable family contained in $\mathscr{N}_x \cup \mathscr{N}^y$. Then $\nu \times \nu A = 0$.*

PROOF (cf. Simms [1989], Lemma 2). Suppose not, then we would have $\nu \times \nu A = 1$. We may suppose $A \subseteq X \cup Y$, where $X \in \mathscr{N}_x$ and $Y \in \mathscr{N}^y$. Hence for all x, $A_x \subseteq X_x \cup Y_x$. By Fubini's theorem applied to $\nu \times \nu$, for almost all x, A_x is full.

Since for all x, X_x is null, it follows that Y_x must be full for almost all x, which is impossible by the previous lemma. □

We can now define the extension μ of $\nu \times \nu$. Let $\mathscr{B}$ be the σ-algebra generated by $\mathscr{A} \times \mathscr{A} \cup (\mathscr{N}_x \cup N^y)$. It is easy to show that $\mathscr{B} = \{B \mid B = A \Delta N,\ A \in \mathscr{A},\ N \in \mathscr{N}_x \cup \mathscr{N}^y\}$. Define μ on $\mathscr{B}$ by $\mu(A \Delta N) = \nu \times \nu A$. μ is a well-defined σ-additive

measure because $A_1 \Delta N_1 = A_2 \Delta N_2$ implies $A_1 \Delta A_2 = N_1 \Delta N_2$; now apply the previous lemma. Obviously μ satisfies $\mu N = 0$ for all $N \in \mathcal{N}_x \cup \mathcal{N}^y$. Furthermore, μ is converse invariant in the sense that $\mu\{\langle x,y\rangle \mid \langle x,y\rangle \in B\} = \mu\{\langle y,x\rangle \mid \langle x,y\rangle \in B\}$.

We can explicitly define a disintegration $\{\mu_x \mid x \in 2^\omega\}$ of μ as follows: put $\mu_x(A_x \Delta N_x) := \nu A_x$. Then μ_x is a σ-additive measure on 2^ω that extends ν, for fixed $B \in \mathscr{B}$ the function $x \to \mu_x B_x$ is ν-measurable and we have

$$(*) \qquad \mu(A \Delta N) = \nu \times \nu A = \int \nu A_x \, d\nu(x) = \int \mu_x(A_x \Delta N_x)\, d\nu(x).$$

This property will serve to validate Q6. We interpret Q in ZFC + $A_{\aleph_0}$ as follows:

(1) $Qx\psi(x) \Leftrightarrow \nu\{x \mid \psi(x)\} = 1$

(2) $Qy\phi(x,y) \Leftrightarrow \mu_x\{y \mid \phi(x,y)\} = 1$, $Qx\phi(x,y) \Leftrightarrow \mu_y\{x \mid \phi(x,y)\} = 1$

for 2-elementary ϕ and ψ. This interpretation is correct because, if ϕ is 2-elementary, either $\forall x\{y \mid \phi(x,y)\}$ is null or $\forall x\{y \mid \neg\phi(x,y)\}$ is null; whence by construction of μ, $\{\langle x,y\rangle \mid \phi(x,y)\}$ and its converse are μ-measurable.

We now show that the Q axioms are valid under this interpretation. Q1–Q3, 5 hold trivially.

We prove Q6 by means of (*):

$$QxQy\phi(x,y) \Leftrightarrow \nu\{x \mid \mu_x\{y \mid \phi(x,y)\} = 1\} = 1 \Leftrightarrow \int \mu_x\{y \mid \phi(x,y)\}\, d\nu(x) = 1$$

$$\Leftrightarrow \mu\{\langle x,y\rangle \mid \phi(x,y)\} = 1 \Leftrightarrow \text{(by converse invariance)} \int \mu_y\{x \mid \phi(x,y)\}\, d\nu(y)$$

$$= 1 \Leftrightarrow \nu\{y \mid \mu_y\{x \mid \phi(x,y)\} = 1\} = 1 \Leftrightarrow QyQx\phi(x,y).$$

The validity of Q4 is due to some simple symmetry properties. Suppose $Qx\phi(x,y)$ and $\forall x(\phi(x,y) \to \psi(x,z))$. Then $\mu_y\{x \mid \phi(x,y)\} = 1$, hence $\mu_y\{x \mid \psi(x,z)\} = 1$. Let π: $2^\omega \to 2^\omega$ be an automorphism that maps y to z and let $\tau = \mathrm{id} \times \pi$ be the induced automorphism on 2^ω. Obviously μ is invariant under τ. Hence if we put $C := \{\langle x,z\rangle \mid \psi(x,z)\}$ we have $\mu C = (\mu\tau^{-1})C$ which implies $\mu_z C_z = (\mu\tau^{-1})_z C_z = \mu_y C_z = 1$. □

6.2.10. COROLLARY. *$\mathrm{ZF}\mathscr{E}_2$ + AC is interpretable in* ZFC + $2^{\aleph_0} = \aleph_2$.

Analogous results hold for n iterations of Q, $n \geq 3$.

So the difference between Freiling's approach and ours lies in the choice of parameters in Q-formulas: we consider either real parameters or arbitrary set parameters, Freiling also considers functions which take countable sets as values. However, in both cases the axioms can be taken to be the same.

REFERENCES

M. J. BEESON [1985], ***Foundations of constructive mathematics***, Springer-Verlag, New York.

J. F. A. K. VAN BENTHEM [1990], *Inference, methodology and Semantics*, ***Festschrift for V. A. Smirnov*** (P. Bystrov, editor), Akademiya Nauk USSR, Moscow.

P. BENACERRAF and H. PUTNAM (editors) [1985], ***Philosophy of mathematics: selected readings*** (second edition), Cambridge University Press, Cambridge.

P. BERNAYS [1985], *On Platonism in mathematics*, ***Philosophy of mathematics: selected readings***, second edition (P. Benacerraf and H. Putnam editors), Cambridge University Press, Cambridge, 1985.

G. BOOLOS [1985], *The iterative concept of set*, ***Philosophy of mathematics: selected readings***, second edition (P. Benacerraf and H. Putnam editors), Cambridge University Press, Cambridge, 1985.

E. BOREL [1909], *Les probabilitiés dénombrables et leurs applications arithmétiques*, reprinted in ***Leçons sur la théorie des fonctions***, Gauthier-Villars, Paris (1914).

P. M. COHN [1981], ***Universal algebra***, Reidel, Dordrecht.

R. O. DAVIES [1963], *Covering the plane with denumerably many curves*, ***Journal of the London Mathematical Society***, vol. 38, pp. 433–438.

R. DIACONESCU [1975], *Axiom of choice and complementation*, ***Proceedings of the American Mathematical Society***, vol. 51, pp. 176–178.

H. FRIEDMAN [1973], *The consistency of classical set theory relative to a set theory with intuitionistic logic*, this JOURNAL, vol. 38, pp. 315–319.

C. FREILING [1986], *Axioms of symmetry: throwing darts at the real number line*, this JOURNAL, vol. 51, pp. 190–200.

K. GÖDEL [1990], ***Collected works***, vol. II (S. Feferman et al., editors), Oxford University Press, London, New York.

N. D. GOODMAN and J. MYHILL [1978], *Choice implies excluded middle*, ***Zeitschrift für Mathematische Logik und Grundlagen der Mathematik***, vol. 24, p. 461.

T. SLAMAN and T. TUGUÉ (editors) [1989], *Mathematical logic and applications*, Lecture Notes in Mathematics, vol. 1388, Springer-Verlag, Berlin.

M. KAUFMANN [1983], *Set theory with a filter quantifier*, this JOURNAL, vol. 48, pp. 263–287.

K. KUNEN [1980], ***Set theory***, North-Holland, Amsterdam.

G. KREISEL [1969], *Two notes on the foundations of set theory*, ***Dialectica***, vol. 23, pp. 93–114.

A. H. KRUSE [1967], *Some notions of random sequence and their set theoretic foundations*, ***Zeitschrift für Mathematische Logik und Grundlagen der Mathematik***, vol. 13, pp. 299–322.

P. MADDY [1988], *Believing the axioms* I, this JOURNAL, vol. 53, 481–511.

M. VAN LAMBALGEN [1990], *The axiomatization of randomness*, this JOURNAL, vol. 55, pp. 1143–1167.

———, [1991], *Natural deduction for generalized quantifiers*, ***Quantification in the Netherlands*** (van der Does and van Eyck, editors), Institute for Logic, Language and Information, University of Amsterdam, Amsterdam.

J. C. OXTOBY [1980], ***Measure and Category***, Springer-Verlag, Berlin and New York.

J. C. OXTOBY and S. KAKUTANI [1950], *Construction of a non-separable invariant extension of the Lebesgue measure space*, ***Annals of Mathematics***, vol. 52, pp. 580–590.

J. SIMMS [1989], *Traditional Cavalieri principles applied to the modern notion of area*, ***Journal of Philosophical Logic***, vol. 18, pp. 275–314.

S. SHELAH [1984], *Can you take Solovay's inaccessible away*?, ***Israel Journal of Mathematics***, vol. 48, pp. 1–47.

R. M. SOLOVAY [1970], *A model of set theory in which every set of reals is Lebesgue measurable*, ***Annals of Mathematics***, vol. 92, pp. 1–56.

J. STERN [1985], *Regularity properties of definable sets of reals*, ***Annals of Mathematical Logic***, vol. 29, pp. 289–324.

A. S. TROELSTRA [1977], ***Choice sequences***, Oxford University Press, London and New York.

A. S. TROELSTRA and D. VAN DALEN [1988], ***Constructivism in mathematics***, vol. II, North-Holland, Amsterdam.

J. VAN AKEN [1986], *Axioms for the set theoretic hierarchy*, this JOURNAL, vol. 51, pp. 992–1004.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF AMSTERDAM
AMSTERDAM, THE NETHERLANDS

E-mail: michiell@fwi.uva.nl