CHAPTER XIV

Algorithms and consistency

14.1. Introduction.

The purpose of this chapter is to present results on decision procedures and computability which are at variance with current presentations of mathematics, although a careful reading of some leading texts shows that, strictly speaking, we are in agreement with some relevant issues in them. To put it boldly, there are no statements which cannot be decided within the natural number system. Indeed in chapter XI we gave algorithms to solve all polynomial equations with complex roots by convergent matrix methods. In *Number, space and logic* [Ad18] we will extend the results of this chapter to include more general logics, called colour logics, we will introduce a proof system first developed by Gentzen called sequent calculus and we also extend Gentzen's methods to prove the consistency of analysis.

We have already discussed four ways in which our discussion of set and number theory is different from some standard interpretations. This is firstly the set theory mZFC, in which the axiom of restricted comprehension in ZFC (Zermelo-Fraenkel set theory with the axiom of choice), which allows only sets satisfying a true predicate, is replaced in mZFC by the axiom of extended comprehension, which also allows the void set satisfying an everywhere false predicate. Second is our discussion previously developed in chapter VII, of the inconsistency of the countability of the rational numbers $\mathbb{Q}$, together with the axiom of induction for natural numbers, in comparison with the uncountable continuum hypothesis of set theory. Third is our extension of nonstandard analysis to ladder algebra. Nonstandard analysis is usually described in terms of ultrafilters. We have shown that a model of this exists for infinitesimals, and that there is an extension of the model as the algebra of ladder numbers. Fourthly, our view on real numbers is that they can be represented by series of uncountable length.

In this chapter we provide an introduction to formal languages in which results in set and number theory may be precisely expressed.

On number theory, we introduce the system Z_1 of elementary arithmetic, its superexponential analogue X_1 and the set theoretic arithmetic, Z_2. We demonstrate their interrelationships and equivalence, which shows that noncommutative and nonassociative logic can be implemented in standard set theory.

We study primitive recursive (PR) and general recursive (GR) functions, for which there is an effective method of computation, and discuss some other number systems. On GR functions, which are described by systems of equations, we point out that a function $f(x) = 0$ is GR and $f(x) = 1$ is GR, but when one of these is adjoined to the other, the result is not GR because it is inconsistent, so that the 'diagonal argument' given by Gödel on the undecidability of problems in Z_2 does not work.

Zermelo-Fraenkel set theory is developed in more detail than we presented in chapter III, and we discuss our nonstandard results on ordinals and cardinals in this context. In *Number, space and logic* this set theory will be extended to a discussion of colour sets, multivalued sets where the two values true and false are replaced by a number of values.

We introduce a discussion of well-ordering, where an ordered set has a lowest member, then ordinal and cardinal infinities. We deviate from current understandings in saying that all

cardinal numbers in systems derived from the natural numbers are countable, that is, we are claiming that the uncountable continuum hypothesis for such sets is false.

We present Gödel's completeness theorem on the consistency of the rules of the propositional calculus. There seems to be a contradiction in the literature between Gentzen's completeness theorem, which uses infinite proofs to show that the axioms of arithmetic are consistent, and Godel's first and second incompleteness theorems, that an inspected system cannot be proved consistent within itself. The incompleteness theorems can be derived by an extension of the 'diagonal argument' method used in our discussion of the existence of undecidable problems in arithmetic, which we have deconstructed, and which we enlarge to these incompleteness theorems in section 8.

14.2. Formal language. [5Co66]

In 1908, Zermelo presented a formal set of axioms for set theory which covered all reasoning in mathematics. The axiomatisation of set theory was in keeping with the spirit of the school of Formalism, led by David Hilbert. In the Formalist point of view, mathematics was seen as a purely formal game played with marks on paper, where the game must not lead to an inconsistent result.

We have met in chapter III, section 3, Peano's axioms for the natural numbers:

(i) Each natural number n has a unique successor n + 1.
(ii) There is a natural number 0 which is not the successor of any natural number.
(iii) Two distinct natural numbers cannot have the same successor.
(iv) If $\mathbb{N}_{U0}$ is a set of numbers so that 0 is in $\mathbb{N}_{U0}$, and such that if a number is in $\mathbb{N}_{U0}$ then its successor is in $\mathbb{N}_{U0}$, then every natural number is in $\mathbb{N}_{U0}$.

These axioms cannot be expressed in a form suitable for a computing machine. This is because the axiom of induction speaks about *sets* of integers, but the axioms do not give rules for assembling sets nor other basic properties of sets. In chapter VII, section 2, however we did provide a rule for induction when the axioms for a set are given.

The symbols used in a formal language, which are not the minimal set of symbols in the interpretation system we will give, we will designate as follows

NOT	&	OR	⇒	⇔
not	and	or	implies	if and only if
∀	∃	=	(,)	x, ′
for every	there exists	equals	parentheses	variable symbols

In principle, the names used under these symbols which give their meaning are independent of the formal language.

We have already met the symbols in the first row in chapter XII. They are connectives used in a logic called the *propositional calculus*. The symbols used in the combined first and second rows are the symbols of the *predicate calculus*. On the second row the symbol ∀, for every (or for all), is known as the universal quantifier, and the symbol ∃, there exists (or for some), is the existential quantifier. Parentheses are used in the language to ensure its precise evaluation. When this evaluation is obvious, they may be omitted. The variable symbols will be formed as x, x′, x″, etc., and we will often rewrite them as x_1, x_2, x_3, etc. if that is useful.

From this set of symbols we can create others. Variable symbols with a special defined use in a system, such as the identity element for a group, are called constants, and will be denoted by c, c′, etc. The variable symbols with the parenthesis symbol can be used to describe *relation symbols*, R_1, R_2, R_3, etc. Each R_k is provided with a natural number n_k which gives the number of variables in the parentheses which follow. To give examples of how these symbols could be used, for $n_0 = 2$ we can define a binary predicate $R_0(x_1, x_2)$, and this could represent $x_1 \leq x_2$, for $n_1 = 3$ we can define a ternary predicate $R_1(x_1, x_2, x_3)$, and this could represent $x_1 + x_2 = x_3$.

By these means symbols in the formal language can be used to express relations in arithmetic. For example if we wanted to express the uniqueness of addition in our language, this is

$$\forall x\ \forall y\ \forall z\ \forall u\ ((R_1(x, y, z)\ \&\ R_1(x, y, u)) \Rightarrow z = u.$$

Similarly, if $R_2(x, y, z)$ represents multiplication as $x.y = z$, then we can express associativity of multiplication in the formal language by

$$\forall x\ \forall y\ \forall z\ \exists t\ \exists u\ \exists v\ ((R_2(x, y, t)\ \&\ R_2(t, z, u)\ \&\ R_2(y, z, v)\ \&\ R_2(x, v, u)).$$

Although the existence of the additive identity zero can be expressed as $\exists x\ \forall y\ (R_1(x, y, y))$, rather than restate this every time we use it, we will introduce the axiom $\forall x\ (x + 0 = x)$. We can now see a way of, say, describing the axioms of a field given in chapter III, section 4, by using formal language.

We now give the rules for writing grammatically correct statements in a formal language. These statements are called *well-formed formulas* (*wff 's*).

(1) Let x and y be variables, c and c′ constants. Then $x = y$, $x = c$ and $c = c'$ are wff 's.
(2) If R is a relation symbol and $u_1, \ldots, u_n$ constants or variables, then $R(u_1, \ldots, u_n)$ is a wff.
(3) If A and B are wff 's then so are NOT (A), (A) & (B), (A) OR (B), (A) $\Rightarrow$ (B) and (A) $\Leftrightarrow$ (B).
(4) If A is a wff, so are $\exists x$ A and $\forall x$ A.

The set of *subformulas* of a wff A is defined as the smallest set which contains A and is closed under the following rules

(i) If NOT (B) is a subformula of A, so is B.
(ii) If (B) & (C), (B) OR (C) or (B) $\Rightarrow$ (C) are subformulas of A, so are B and C.
(iii) If either $\exists x$ B or $\forall x$ B are subformulas of A, so is B(c) for a constant c.

In subsequent developments we will need to distinguish between the ideas of *free* and *bound* variables. An example is as follows. Suppose we had $\forall x\ y = z$, which by rule (4) is a wff. Now $y = z$ does not involve x, so the quantifier $\forall x$ has no effect. A similar circumstance might be where variable and constant symbols occur in a wff without referencing anything.

Definition 14.2.1. A variable symbol in a wff is *free* or *bound* according as

(a) Every variable occurring in a formula for rules (1) and (2) above is free.
(b) Free and bound occurrences derived from rule (3) above are the same as those for A and B.
(c) Free and bound occurrences of variables in $\exists x$ A and $\forall x$ A are the same as those for A except free occurrences of x are now called bound.

We allow binding to occur more than once. If this does not occur, the wff is called *good*.

Definition 14.2.2. A *statement* is a formula with no free variables.

Let us recall the rules for the propositional calculus given in chapter XIII. We introduced two variables, τ, which we will now identify with true, and υ, which we will identify with false.

Then the truth tables for statements A and B map A(τ, υ), B(τ, υ) to NOT A, A & B, A OR B, A ⇒ B and A ⇔ B as follows

NOT A	A & B	A OR B	A ⇒ B	A ⇔ B	A	B
υ	τ	τ	τ	τ	τ	τ
υ	υ	τ	υ	υ	τ	υ
τ	υ	τ	τ	υ	υ	τ
τ	υ	υ	τ	τ	υ	υ

The number of possible truth table functions is four for a function of one variable, 16 for a function of two variables, and in general 2↑(2↑m) for a function of m variables.

In fact, we can reduce the number of tables required to just NOT A and, say, A & B, since

A OR B = NOT(NOT A & NOT B)
A ⇒ B = NOT(A & NOT B)
A ⇔ B = (A ⇒ B) & (B ⇒ A).

For the predicate calculus we can also introduce a simplification, since ∀ can be replaced by NOT ∃ NOT.

The objective of mathematics is to discover true theorems. The intention is that any valid statement is intuitively true in any interpretation of the relation and constant symbols used to form it. We use the term *valid* to describe statements formed by the rules we give next.

(A) *Rule of the propositional calculus*. If P is a propositional function in the variables A_1, ..., A_n that is always true, then on replacing any A_k by a valid statement results in a valid statement.

(B) *Rule of inference*. This allows us to form new valid statements from old ones. If A and (A) ⇒ (B) are valid statements, so is B.

(C) *Rules of equality*. These allow us to manipulate equal signs. Let c, c′ and c′′ be constants.

(i) c = c′, (c = c′) ⇒ (c′ = c) and ((c = c′) ⇒ (c′ = c′′)) ⇒ (c = c′′) are valid statements.

(ii) If A is a statement, and if A′ represents A with every occurrence of c replaced by c′, then (c = c′) ⇒ ((A) ⇒ (A′)) is a valid statement.

(D) *Change of variables*. If A is a statement and A′ is obtained from A by replacing each occurrence of the symbol x by x′, then (A) ⇔ (A′) is a valid statement. This also applies to good statements, as we have defined them.

(E) *Rule of specialisation*. Let A(x) be a formula with one free variable x in which every occurrence of x is free, and A(c) be result from replacing every occurrence of x by the constant symbol c. Then (∀x A(x) ⇒ A(c)) is a valid statement.

(F) Let B represent a statement where c and x are absent. If A(c) ⇒ B is valid, so is ∃x A(x) ⇒ B. This allows us to argue about a statement using an ‘arbitrary’ constant c, and infer that ∀x A(x), since we have used no special properties of c.

(G) *Rearrangement of quantifiers to the beginning of a statement*. Let A(x) have x as its only free variable, where every occurrence of x is free, and B be a statement with x absent.

(i) (NOT (∀x A(x))) ⇔ (∃x NOT (A(x))).

(ii) ((∀x A(x)) & (B)) ⇔ (∀x ((A(x) & B))).

(iii) ((∃x A(x)) & (B)) ⇔ (∃x (A(x)) & (B)).

14.3. Number theory and primitive recursive functions. [5Co66]

We will describe the axioms for a formalisation of elementary number theory denoted by Z_1, in which statements are of relatively simple type, using only + and × (we denote × also by .), and the corresponding superexponential extension of it, which we denote by X_1, which extends the rules for + and × to the superexponential operators $^n|$ which are obtained by general operations nested with parentheses on the left, and $|^n$ operators obtained from operations nested on the right. This is described in chapter XVII. We will compare Z_1 with another formalisation of elementary number theory, Z_2, where the elements are finite sets. We will show that the axioms for X_1 imply those for Z_1, and that we can construct the axiom system X_1 by an interpretation of Z_1. Further, we will show Z_1 holds if and only if Z_2 holds; we will develop the idea of primitive recursive functions in order to prove this theorem.

The axiom system for Z_1 has non-negative elements. We introduce two ternary relations, the addition relation $R_1(x, y, z)$ and $R_2(x, y, z)$ for multiplication, but we use $x + y = z$ for R_1 and $x.y = z$ for R_2 to be read more easily. There are two constant symbols, 0 and 1. From now on we use the notation $\exists!x\, A(x)$ meaning "there exists a unique x ... " for $\exists x\, \forall y(A(y) \Leftrightarrow x = y)$.

Axioms for Z_1.

(1) $\forall x, y\ \exists!z\ (x + y = z)$.
(2) $\forall x, y\ \exists!z\ (x.y = z)$.
(3) $\forall x\ (x + 0 = x)\ \&\ (x.1 = x)$.
(4) $\forall x, y\ (x + (y + 1) = ((x + y) + 1)$.
(5) $\forall x, y\ (x.(y + 1) = x.y + x)$.
(6) $\forall x, y\ (x + 1 = y + 1 \Rightarrow x = y)$.
(7) $\forall x\ (\text{NOT } x + 1 = 0)$.
(8_n) $\forall u_1, ..., u_k\ [A_n(0, u_1, ..., u_k)\ \&\ \forall y\ (A_n(y, u_1, ..., u_k) \Rightarrow (A_n(y + 1, u_1, ..., u_k)))$
$\Rightarrow \forall x\ A_n(x, u_1, ..., u_k)]$.

Axioms (8_n) for Z_1 are what is properly called an axiom scheme, that is, we have listed the countably many formulas with at least one free variable x, of the form $A_n(x, u_1, ..., u_k)$, where k is dependent on n.

We design the axioms for the superexponential system X_1 by analogy with those for Z_1. These axioms are divided into parts for left nested operators and those for the right. The constant symbols for left nest operators are now v_n where v_n is the left neutral element of chapter XVII, section 8, with w_n denoting a right neutral element. There are induction axiom schemes for left nest operators (f_n) and (g_n), and for right nest operators to follow.

Axioms for left nest operators in X_1.

(a) $\forall x, y, n\ \exists!z\ (x\ ^n|\ y = z)$.
(b) $\forall x\ \exists!v_n\ \forall n\ (v_n\ ^n|\ x = x)$.
(c) $\forall x\ \exists!w_n\ \forall n\ (x\ ^n|\ w_n = x)$.
(d) $\forall x, m, n\ (\text{NOT } x = v_n)\ \&\ (\text{NOT } v_n\ ^n|\ x = v_m)\ \&\ (\text{NOT } n = m)$.
(e) $\forall x, m, n\ (\text{NOT } x = w_n)\ \&\ (\text{NOT } x\ ^n|\ w_n = w_m)\ \&\ (\text{NOT } n = m)$.
(f_n) $\forall u_1, ..., u_k\ [A_n(v_n, u_1, ..., u_k)\ \&\ \forall y\ (A_n(y, u_1, ..., u_k) \Rightarrow (A_n(v_{n+1}\ ^{n+1}|\ y, u_1, ..., u_k)))$
$\Rightarrow \forall x\ A_n(x, u_1, ..., u_k)]$.
(g_n) $\forall u_1, ..., u_k\ [A_n(w_n, u_1, ..., u_k)\ \&\ \forall y\ (A_n(y, u_1, ..., u_k) \Rightarrow (A_n(y\ ^{n+1}|\ w_{n+1}\ u_1, ..., u_k)))$
$\Rightarrow \forall x\ A_n(x, u_1, ..., u_k)]$.

The axioms for right nest operators are obtained from those above for left nest operators on substituting operators denoted by, say, $^{n}|$, with $|^{n}$ operators. However, as we will see in chapter XVII, section 6, there are rules for $^{n}|$ which are not the same as for $|^{n}$ operators:

(h) $\forall x, y, z, n\ (n > 2) \ \& \ (x\ ^{3}|\ y)\ ^{n}|\ z = x\ ^{3}|\ (y\ ^{n-1}|\ z)$,

but

(i) $\exists x, y, z\ (x\ |^{3}\ y)\ |^{4}\ z \neq x\ |^{3}\ (y\ |^{3}\ z)$.

We also have the following rule for left nest operators

(j) $\forall x, y, n\ x\ ^{n}|\ (y + 1) = (x\ ^{n}|\ y)\ ^{n-1}|\ (x\ ^{n}|\ w_n)$,

where for example this satisfies the following step-down equation

(k) $\forall x, y\ x^{y+1} = x^{y}.x$.

The corresponding step-down equation for right nest operators is

(l) $\forall x, y, n\ x\ |^{n}\ (y + 1) = (x\ |^{n}\ y)\ |^{n-2}\ (\hat{x}\ |^{n}\ (y + 1))$,

where we interpret the final term reducing operator $\hat{x}$ as for instance with $n = 4$, $y = 2$ by

(m) $\forall x\ x\uparrow(x\uparrow x) = x\uparrow(x\uparrow(1 + (x - 1))$
$= (x\uparrow x)(x\uparrow(x\uparrow(x - 1))$.

As a consequence of the axioms for X_1, we can make the following statements

(i) $\forall x, y, n\ v_n\ ^{n}|\ (x\ ^{n}|\ y) = (v_n\ ^{n}|\ x)\ ^{n}|\ y$.
(ii) $\forall x, y, n\ (x\ ^{n}|\ y)\ ^{n}|\ w_n = x\ ^{n}|\ (y\ ^{n}|\ w_n)$.
(iii) $\forall x, y, m, n\ (v_n\ ^{n}|\ x)\ ^{m}|\ y = v_n\ ^{n}|\ (x\ ^{m}|\ y)$.
(iv) $\forall x, y, m, n\ x\ ^{n}|\ (y\ ^{m}|\ w_m) = (x\ ^{n}|\ y)\ ^{m}|\ w_m$.
(v) $\forall x, y, n\ (v_n\ ^{n}|\ x) = (v_n\ ^{n}|\ y) \Rightarrow x = y$.
(vi) $\forall x, y, n\ (x\ ^{n}|\ w_n) = (y\ ^{n}|\ w_n) \Rightarrow x = y$. □

Theorem 14.3.1. *The axioms of* X_1 *imply those of* Z_1*, and there is an interpretation of* Z_1 *in which* X_1 *holds*.

Proof. We note that the statements in X_1 imply those in Z_1 in the case where $^{1}|$ or $|^{1}$ is + and $^{2}|$ or $|^{2}$ is $\times$; for example v_1 and w_1 are 0, and v_2 and w_2 are 1.

Further, Z_1 can be used to define the $^{n}|$ and $|^{n}$ operators in X_1 using the induction scheme (8_n) for Z_1. In particular we can apply this twice, so for double recursion we note that

$\forall u_0, ..., u_k\ [A_n(0, u_1, ..., u_k) \ \& \ (\forall y_0\ (A_n(y_0, u_1, ..., u_k) \Rightarrow (A_n(y_0 + 1, u_1, ..., u_k))))$
$\& \ A_n(0, 0, u_2, ..., u_k) \ \& \ (\forall y_1\ (A_n(0, y_1, u_2, ..., u_k) \Rightarrow (A_n(0, y_1 + 1, u_2, ..., u_k))))$
$\Rightarrow \forall x_0, x_1\ A_n(x_0, x_1, u_2, ..., u_k)]$.

If $A(0, x, y, z, m)$ is axioms (a) & (b) & (c) & (d) & (e) & (j) & (k) for X_1 with $n = 3 = p + 3$, then (8_n) can be expressed as

$\forall x, y, z, m\ [A(0, x, y, z, m) \ \& \ \forall p\ (A(p, x, y, z, m) \Rightarrow (A(p + 1, x, y, z, m)))$
$\Rightarrow \forall q\ A(q, x, y, z, m)]$,

but if axioms (1) to (7) hold for Z_1, then these can be interpreted in X_1, and given these interpretations axiom (j) relates $^{n}|$ to $^{n-1}|$, so by (c) $^{n}|$ is defined by induction scheme (8_n) operating from w_n to y inclusive for all n. Also axiom (j) steps down y, and it eventually reaches w_n, which is precisely what we need to prove. □

Corollary 14.3.2. *The corresponding results to* 14.3.1 *hold with each* $^{i}|$ *replaced by* $|^{i}$.

We are dealing with axiom (l) here, in a similar way to axiom (j) for $^{i}|$. □

In the system for Z_2 that we present next, the elements are finite sets, and are very similar to the axioms we will give later in section 4 which allow infinite sets. The induction axiom in Z_2 is again an infinite axiom scheme.

Axioms for Z_2.

(1) *Extensionality*. $\forall x, y\ (x = y \Leftrightarrow \forall z\ (z \in x \Leftrightarrow z \in y))$. This asserts that a set is determined by its members. Thus in this system there are no so-called atomic sets which contain no members other than the empty set.

(2) *Empty set*. $\forall x$ (NOT $x \in \emptyset$).

(3) *Set of unordered pairs*. $\forall x, y\ \exists z\ \forall w\ (w \in z \Leftrightarrow w = x$ OR $w = y)$.

(4) *Union of sets*. $\forall x, y\ \exists z\ \forall w\ (w \in z \Leftrightarrow w \in x$ OR $w \in y)$. From this we can define an ordered pair as $<x, y>$ as $\{\{x\}, \{x, y\}\}$. We can now define the natural numbers called $\mathbb{N}_{U0}$, $\{0, 1, 2, 3, ...\}$, as the set $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, ...\}$.

The usual model for axioms (1) to (4) above is that of all finite sets which can be built up from $\emptyset$ by a finite number of steps, but we cannot formalise the notion of a natural number in Z_2 until we have an axiom of induction. This motivates the next two definitions.

x *is a natural number if*

(i) $\forall y, z\ (y \in x\ \&\ z \in x \Rightarrow y = z$ OR $y \in z$ OR $z \in y)$.

(ii) $\forall y, z\ (y \in x\ \&\ z \in x \Rightarrow z \in x)$.

The reason this gives the set $\mathbb{N}_{U0}$, defined in terms of 0 as $\emptyset$, 1 as $\{\emptyset\}$, etc. is that, firstly, it defines a minimal element, $\emptyset$, with respect to the $\in$ relation, because by (i) if this is y, every non-empty set x which contains it satisfies $z \in x \Rightarrow$ NOT $z \in y$. Then by (i) and (ii) for a new value of y given by the non-empty set x not including $\emptyset$, its value must be $\{\emptyset\}$, so that the natural numbers are generated.

If x *is a natural number, denote* $x + 1$ *by* $x \cup \{x\}$.

We will show this is a reasonable assignment. We need to show from the previous definition that $x + 1$ is a natural number. If y and z belong to $x + 1$, then there are three alternatives: both belong to x, both belong to $\{x\}$, or one belongs to x and the other to $\{x\}$. Axiom (i) holds because in the first case since both belong to x and x is a natural number, (i) holds, in the second case if a set belongs to $\{x\}$ it must equal x, so that $y = z$, and in the third we have trivially either $y \in z$ or $z \in y$. Also axiom (ii) holds because if $y \in x + 1$ and $z \in y$, if $y = x$ then $z \in x$, so $z \in x + 1$, whereas if $y \in x$ then again $z \in x$, and since x is a natural number, $z \in x + 1$.

Using $A_n(x, u_1, ... u_k)$ with codomain all formulas with at least one free variable, writing 0 for $\emptyset$ and Nat x as an abbreviation for the definition that x is a natural number, we can now formulate

(5_n) *Induction principle*. $\forall u_1, ..., u_k\ [A_n(0,u_1, ..., u_k)\ \&\ \forall y(\text{Nat } y\ \&\ A_n(y, u_1, ..., u_k) \Rightarrow (A_n(y + 1, u_1, ..., u_k))) \Rightarrow \forall x\ (\text{Nat } x \Rightarrow A_n(x, u_1, ..., u_k))]$.

This axiom scheme is equivalent to the axiom scheme for Z_1. Not every possible set of natural numbers is described above by a finite collection of properties A_n for suitable u_i. Further, we have seen that the Peano axioms specify that the set $\mathbb{N}_{U0}$ is unique, but we will construct ladder natural numbers later in this section which satisfy the Peano axioms except uniqueness in Z_2, so we have not captured completely the informal strength of the Peano axioms.

To develop conventional number theory in Z_2, we proceed as follows. To describe a function, this is a set of ordered pairs in which its leftmost member occurs at most once, the domain of the function being the set of these left members, and the codomain the set of right members. A function in elementary number theory is usually described by a constraint condition C(x), so if C(x) can be expressed in Z_2, where x is a set of u ordered pairs {<1, f(1)>, ... <u, f(u)>} and a recursion relation holds amongst the f(i), then this recursion relation can be defined in Z_2. In this way we avoid using an infinite set, and can use the induction principle (5_n).

We now discuss primitive recursive functions, which are a description of functions often occurring in practice. It will be our task to find out how this idea can be used in the formal system Z_2.

A function $f(u_1, ..., u_k)$ from natural numbers $\mathbb{N}_{U0}$ to $\mathbb{N}_{U0}$ is *primitive recursive* (PR) if it is constructed according to the following rules.

(1) c for some constant c is PR
(2) $f(u_1, ..., u_k) = u_i$ for $1 \leq i \leq k$ is PR
(3) $f(u) = u + 1$ is PR
(4) If $f(u_1, ..., u_k)$ and $g_1, ..., g_k$ are PR, so is $f(g_1, ..., g_k)$
(5) If $f(0, u_2, ..., u_k)$ is PR and $g(v, u_1, ..., u_k)$ is PR
and $f(u + 1, u_2, ..., u_k) = g(f(u, u_2, ..., u_k), u, u_2, ..., u_k)$ is PR then f is PR.

Most elementary functions of the natural numbers are PR; addition, multiplication, powers, the factorial function and the nth prime are PR. Because PR functions are defined only for nonnegative numbers, subtraction must be defined for them as subtraction when the result is nonnegative, and zero otherwise. To include truth values in PR functions, define true as the number 1. For false the standard allocation is 0, but elsewhere we may use -1 for this. The PR functions are important because they are effectively computable, and should be thought of as the simplest class of computable functions. Once we have shown that these PR functions can be expressed in a system, we can describe many non-constructive functions by using the propositional connectives NOT, &, OR, etc., for instance the function 'the lowest number for which the Goldbach conjecture is false, if there is one, otherwise 0'.

Reference [Od89] states that a function which is not PR is Ackermann's function:

$A(0, y) = y + 1,$
$A(x + 1, 0) = A(x, 1),$
$A(x + 1, y + 1) = A(x, A(x + 1, y)).$

It can be shown that

$A(0, y) = y + 1,$
$A(1, y) = y + 2 = 2 + (y + 3) - 3,$
$A(2, y) = 2y + 3 = 2(y + 3) - 3,$
$A(3, y) = 2 \uparrow (y + 3) - 3,$

and

$A(x, y) = (2 \,|^x\, (y + 3)) - 3.$

It would follow for finite n that $|^n$ is not PR, but we have shown that the argument for $^n|$ may be carried over to $|^n$, so because we can demonstrate that Z_1 is PR, and since Z_1 holds if and only if an implementation of X_1 holds, where for Z_1 we have applied axiom scheme (8_n) twice to prove this, we have not reached the same conclusions. □

To show that any PR function can be expressed in Z_2, we will prove

Theorem 14.3.3. *If* $f(u_1, ..., u_k)$ *is a* PR *function, there exists a formula* $A(u_1, ..., u_k, v)$ *in* Z_2 *where*

(i) $\forall u_1, ..., u_k \exists! v$ *such that* $A(u_1, ..., u_k, v)$ *is a provable statement in* Z_2

(ii) *If* $f(u_1, ..., u_k) = w$, *then* $A(u_1, ..., u_k, w)$ *is provable in* Z_2.

Proof. The proofs of (1) to (4) are trivial. To prove (5), assume for simplicity that k = 1. Then suppose $g(u_1, u_2)$ is PR and that f is defined by

$f(0) = c$

$f(u + 1) = g(f(u), u)$.

The proof is by induction in which we increase the complexity of the definition of f, so we assume that there is a formula $B(u_1, u_2, v)$ such that $g(u_1, u_2) = v$ if and only if $B(u_1, u_2, v)$ is provable in Z_2. Then $f(u) = v$ is represented by the formula

$\exists f$ (f is a function) & (the domain of f = $(x: x < u)$) & $(f(0) = c)$

& $(f(u) = v)$ & $\forall x (x < u \Rightarrow g(f(u), u) = f(u + 1))$,

so that this formula is a formula $A(u, v)$ in Z_2, and the formula satisfies the conditions of the theorem. □

We will carry out the proof in the reverse direction, that any statement in Z_2 is equivalent to a corresponding statement in Z_1. In the above proof for Z_2 we used the fact that we can use finite sequences of arbitrary length. In Z_1 the corresponding proof is non-trivial. In order to define precisely arbitrary sequences of finite length in Z_1 we need

(i) A proof in Z_1 of the prime number theorem, described in chapter III, section 9, that every natural number up to order of factors is uniquely the product of primes.

(ii) The Euclidean algorithm in Z_1, again described in chapter III, section 9, that every natural number can be written for a natural number u uniquely in the form au + b.

(iii) The Chinese remainder theorem in Z_1, which is described in chapter VI section 2, stating that a system of k simultaneous linearly independent equations with each equation holding (mod u_k), can be reduced to a solvable set of linear equations all taken (mod the l.c.m. of the u_k).

(iv) A proof in Z_1 that the factorial function u! is greater than u for u > 1 a natural number.

We need to prove that if f is a PR function, there is a formula A in Z_1 such that $f(u_1, ..., u_k) = x$ if and only if $A(u_1, ..., u_k, x)$. To do this we will use the results of number theory indicated above, and prove two preliminary lemmas. When we have the theorem, we will go back to the lemmas and prove them directly within Z_1.

To prove the first lemma, we will introduce the formula $x = 1 + (i + 1)d$, which we denote by $A'(d, i, x)$.

Lemma 14.3.4. *There is a formula* $A'(d, i, x)$ *in* Z_1 *for which*

(i) $\forall d, i \exists!\, x\, A'(d, i, x)$.

This means x is a function of d and i, which we will write as x(d, i).

(ii) $\forall u, M \exists d \forall i, j [i < j < n \Rightarrow x(d, i) > M$ & (x(d, i) *and* x(d, j) *are relatively prime*)].

This says that for every u and M, we can find a d with all x(d, i) for i < u relatively prime and greater than M.

We have used abbreviations: x > y can be written as $\exists z$ (NOT z = 0 & x = y + z), whereas "x and y are relatively prime" can be expressed as

$\forall v, w, z[x = v.z \;\&\; y = w.z \Rightarrow z = 1]$.

Proof. If u and M are relatively prime, we can put d = (max(u, M))!, giving x(d, i) > M. Then if i < j < u, if a prime p divides x(d, i) and x(d, j), from the definition it must divide (i – j)d. If we assume p divides d, then it cannot divide x(d, i), but if it divides (i – j), then p < u, so that p divides d, a contradiction with the definition of x(d, i). □

We will now denote the remainder when c is divided by x(d, i) by r(c, d, i).
Lemma 14.3.5. *For any sequence of natural numbers* a_1, ..., a_u, *there is a* c *and* d *for every* i ≤ u *with* a_i = r(c, d, i).

Proof. Note that this lemma as it stands at the moment cannot be directly stated in Z_1, since it refers to sequences of arbitrary length. We will address this issue in lemma 14.3.8. To prove the lemma using generally understood arithmetic, choose d so that for i ≤ u, x(d, i) is greater than max a_i. Then by the Chinese remainder theorem we can find the c we want. □

Theorem 14.3.6. *If* f *is a* PR *function, there is a formula* A *in* Z_1 *such that* $f(u_1, ..., u_k) = x$ *if and only if* $A(u_1, ..., u_k, x)$.

Proof. Once again we will restrict discussion to the nontrivial case of rule (5) for a PR function. Let us assume that $f(0, u_2, ..., u_k)$ and g are expressible in Z_1, with $A(u_1, ..., u_k, x)$ the formula

$$\exists c, d\, [f(0, u_2, ..., u_k) = r(c, d, 0) \,\&\, x = r(c, d, u_1) \,\&$$
$$\forall i\, [i < u_1 \Rightarrow r(c, d, i + 1) = g(r(c, d, i), i, u_2, ..., u_k)]],$$

which implies since any sequence a_0, ..., a_{u_1} is in the form r(c, d, i) for some c and d, that we have obtained the required definition of f. □

We now address the issue of expressing lemma 14.3.4 in the format of the system Z_1, so we will need to prove
Lemma 14.3.7. *Let* a, b, u *and* M *be natural numbers and* p *be a prime number*
(i) *If* a *and* b *have a common divisor, they have a prime as a common divisor*.
(ii) *If a prime* p *divides* u = ab, *then* p *divides either* a *or* b.
(iii) *For every* M, *there exists a number* (*for example* M!) *divisible by every* u ≤ M.

Proof. The principle of induction in Z_1 allows us to state that for every property A(u) in Z_1 that there is a least u satisfying A(u). The proof of (i) proceeds from the existence of a least nontrivial common divisor, which must therefore be prime.

To prove (ii) in Z_1 we know if u = 2 then the result is proved. If u is composite it can be represented by the product of numbers which are greater than 1 and less than u, but we know that by the induction method in Z_1 that a and b are either prime or the product of primes.

To prove (iii) by induction in Z_1, if x is divisible by all u ≤ M, then x.(M + 1) is divisible by all u ≤ M + 1. □

We now proceed to formulating lemma 14.3.5 in Z_1.
Lemma 14.3.8. *Let* a, c, d, u *and* i *be natural numbers where the* x(d, i) *are relatively prime for all* i ≤ u + 1. *Then there exists a natural number* c′ *such that* r(c, d, i) = r (c′, d, i) *for* i ≤ u *and* r(c, d, u + 1) = a.

Proof. This statement allows us to extend sequences in Z_1 by one element at a time. But for all i ≤ u there exists a q such that x(d, u + 1) is prime to q, where for j ≤ i, x(d, j) divides q. This means there is a q with x(d, i) dividing q for every i < u, where x(d, u +1) is prime to q. This means that for some v, if c′ = c + qv, c′ satisfies the property of the theorem. □

We can now prove by induction in Z_1 that the given formula f does indeed define a unique PR function, where the properties of a PF function in Z_1, such as composition of functions, holds.

From 14.3.6 we can prove by induction in Z_1 that the formula given to define the PR function f defines a unique f. It is remarkable that + and × are sufficient to define PR functions. Using it, we can show that Z_1 is as powerful as Z_2, in the sense that Z_2 can be embedded in Z_1. Indeed, for natural numbers u_k with $u_0 = 0$, define corresponding sets S_k with $S_0 \neq \emptyset$. We can define a function f(t, u) such that f is 1 if S_t belongs to S_u, and is equal to zero otherwise. Then f is a PR function. We can now prove the axioms of Z_2 in Z_1 if we define $x \in y$ if and only if f(x, y) = 1. It can now be shown that any statement provable in Z_1 is provable in Z_2, which follows from the fact that + and × are PR functions. □

In chapter VII, section 3, we provided a ladder number analogue $\mathbf{L}_{\mathbb{N}U0}$ of the natural numbers, where natural ladder numbers $\mathbf{L}_{\mathbb{N}U0}$ are defined by

$$\mathbf{L}_{\mathbb{N}U0} = \bigcup_m[\mathbb{N}_{U0}(\mathbf{\Omega}_{\mathbb{N}})^m],\ m \in \mathbb{N}_{U0}.$$

We assume, for example, $n \neq n(\mathbf{\Omega}_{\mathbb{N}})^m$. We stress here that we interpret

$$0 \neq 0(\mathbf{\Omega}_{\mathbb{N}})^m,$$

as is also the case for a vector space. Then the operations on ladder natural numbers may be considered as operations on a vector space with basis elements $(\mathbf{\Omega}_{\mathbb{N}})^m$. The algorithmic proof of a proposition by induction: choose a start value, assume for n and then prove for n + 1, now extends to $n \in \mathbf{L}_{\mathbb{N}U0}$. Thus directly, when an induction on n works for $\mathbb{N}_{U0}$, induction on $n'(\mathbf{\Omega}_{\mathbb{N}})^m$ also works for $\mathbf{L}_{\mathbb{N}U0}$, and thus (by induction) on $n + n'(\mathbf{\Omega}_{\mathbb{N}})^m$. Since the Peano axiom of induction states that the natural numbers are unique, we must augment it for $\mathbb{N}_{U0}$ by saying it contains no elements $(\mathbf{\Omega}_{\mathbb{N}})^m$.

Ladder numbers are ordered by their *rungs* $n(\mathbf{\Omega}_{\mathbb{N}})^m$. We may wish to choose a ladder number analogue of the natural numbers for which only the lowest rung has a lowest element. In this case, choose

$$\mathbb{N}_{U0}\bigcup_m[\mathbb{Z}(\mathbf{\Omega}_{\mathbb{N}})^m],\ m \in \mathbb{N}_{\neq 0},$$

where $\mathbb{Z}$ is the set of the positive , negative or zero integers. □

We stated in chapter I, section 2, that there exist nonstandard models of arithmetic which are not trivial and where 1 = 0. There are various such models. Let K_p be the Eudoxus numbers (mod p), where p is a prime. Then this has the operations of + and × defined within it, except that division by zero is not defined.

For rational numbers (mod p), define

$$\frac{j + kp}{m + np} = \frac{j}{m} \text{ (mod p), provided } m \neq 0.$$

For Eudoxus numbers (mod p), take the denominator to be 1.

Now define the map

$$\text{(mod p)} \rightarrow \text{(mod 1)}$$

by taking the non integer part. For division by zero, take this as division by 1. This is an implementation of K_p.

We can also define a K_r algebra, where instead of p we have an algebraic, or more generally a real number r.

It is possible to define a hyperintricate algebra $K_Я$, with components v_B (mod u)B, where B is a hyperintricate basis element. □

14.4. General recursive functions.

In this section we will look at a generalisation of PR functions: the general recursive (GR) functions. A motivation in considering such functions was the *decision problem*: how do we mechanically decide if any problem in a system (say X_1) is true? This was answered in the negative by the use of a diagonal argument for GR functions. An objective of our account is to deconstruct the use of diagonal arguments. We proceed as follows.

The first definition of a GR function was given by Gödel in 1934, building on a suggestion of Herbrand. These GR functions include the PR functions. Before giving this definition, we note that at first there appears to be an obstacle to showing that a GR function, however it is to be defined, may not be effectively computable. Namely, if $f_n(x)$ are a set of GR functions, then on setting $g(n) = f_n(n) + 1$, $g(n)$ would be considered to be recursively defined but not a GR function. The way out of this is to give a non-effective definition of a GR function. We construct this definition from the observation that PR functions satisfy the condition that they may be defined from a finite set of equations involving functions. We proceed by developing a set of conditions that are satisfied by functions, and then define GR functions in terms of the relations between these functions.

Consider an alphabet of function symbols f, g, h, … each associated with a fixed number of variables with values in the natural numbers, denoted by x, y, z, …, and the symbols "0", "=", "′" where x′ means x + 1, with parentheses "(", ")" and the comma ",". Gödel defines

(1) A *numeral* is an expression of the form 0, 0′, 0″, … etc.

(2) A *term* is defined by

- (i) 0 is a term
- (ii) the variables x, y, z are terms
- (iii) if t is a term, so is t′
- (iv) if $t_1, \ldots, t_n$ are terms and f is a function in n variables, then $f(t_1, \ldots, t_n)$ is a term.

(3) An *equation* is an expression in the form t = u, where t and u are terms.

(4) If E is a finite set of equations, a *deduction* from E is obtained from applying the following rules:

Rule 1. Given an equation all occurrences of a variable x may be replaced by a numeral.
Rule 2. If $f(t_1, \ldots, t_n) = u$ has been deduced, with u and t_k numerals, then in any equation occurrences of $f(t_1, \ldots, t_n)$ may be replaced by u.
Rule 3. If t = u can be deduced, so can u = t.

Definition 14.4.1. A function $f(u_1, \ldots, u_n)$ is *general recursive* (GR) if there is a finite set of equations so that for any of the numerals $u_1, \ldots, u_n$ there is a unique v such that $f(u_1, \ldots, u_n) = v$ can be deduced.

We will give examples. The system

$f(t, 0) = t$,
$f(t, u') = f(t, u)'$

defines f = t + u. We can define systems with recursion in more than one variable:

$f(0, t') = t'$
$f(u, 0) = u'$
$f(u', t') = f(u, f(u', t))$

defines a function $f(u_1, u_2)$ under the condition that if $g(u_2)$ is PR then there exists a u_1 satisfying $\forall u_2\ f(u_1, u_2) > g(u_2)$.

A consequence of their definitions is that every PR function is GR.

Church's thesis is the unprovable statement that every effectively computable function is a GR function. We can consider functions defined by an infinite number of equations, which are not GR. For example, let the function on an infinite number of arguments $g(u_1, u_2, u_3, \ldots)$ have codomain the vector $(u_1, u_1 + u_2, u_1 + u_2 + u_3, \ldots)$, but we might not call g effectively computable.

Remark 14.4.2. We note that a program to compute the decimal expansion of $\sqrt{2}$ never terminates, but the value of $\sqrt{2}$ is unique. On the other hand, if a value of F(x) is either 0 or 1, and if at the nth iteration of an algorithm it computes F(x) = 1 when F(x) = 0 and F(x) = 0 when F(x) = 1, then if we assume that functions with unique members of the domain do not map to multiple sets in the codomain, then F(x) is inconsistent and its computed value never terminates. We conclude from the halting of an algorithm that the value of a function has been obtained, but we cannot conclude just from the non-halting of a program whether a function is unique and exists but is not finitely computed, or whether the function has been defined so that it is not unique, and therefore is in contradiction with the definition of a function.

Example 14.4.3. Note that f(u) = 0 is PR and therefore GR, and so is f(u) = 1, but the combination of these equations together is a contradiction (it violates $\forall x$ (NOT x + 1 = 0) and so does not define a unique v with f(u) = v, and is thus not GR. This example is important for the deconstruction of the proof given in [5Co66] which follows.

In more detail, let $i = 1, \ldots, n \in \mathbb{N}$, with k and n depending on i, where k(i) and $n_{i\,k(i)} \in \mathbb{N}$. Let E_i be a set of equations denoted by

$$(f_{i\,k(i)} = n_{i\,k(i)}).$$

The definition of deduction in Cohen is in terms of itself, as a set of rules satisfying rules of deduction.

Definition 14.4.4. By a finite set of n equations E'_n we mean the set

$$E'_n = E_1 \cup E_2 \cup \ldots \cup E_n,$$

where E_i satisfies the set of equations given already and E'_n satisfies

$$(f_{1\,k(1)} = n_{1\,k(1)}) \,\&\, (f_{2\,k(2)} = n_{2\,k(2)}) \,\&\, \ldots \,\&\, (f_{n\,k(n)} = n_{n\,k(n)}).$$

This set of functions is *general recursive* (GR) if and only if there is no deduction

$$f_{i\,k(i)} = n_{i\,k(i)}, \quad f_{i\,k(i)} = m_{i\,k(i)}$$

with

$$n_{i\,k(i)} \neq m_{i\,k(i)}.$$

Assume E'_n is GR. Consider the set E'_1 of an equation satisfying for a new function

$$f'_{11} = n'_{11}$$

where

$$n'_{11} \neq n_{i\,k(i)}$$

for any i. Then this set of functions is GR, but the set of functions

$$E'_{n+1} = E'_n \cup E'_1$$

is not GR by definition. □

Theorem 14.4.5. *A function is* GR *if and only if there exists a set of functions which are* GR.

Proof. Let $f(u_1, \dots, u_n)$ be a function. If $f(u_1, \dots, u_n)$ is GR there is both a unique v such that $f(u_1, \dots, u_n) = v$ and this equation belongs to a set (given by itself) which is GR. If f is not GR there is no unique v such that $f(u_1, \dots, u_n)$ can be deduced. Thus there exist the set of equations $f(u_1, \dots, u_n) = v_1$ and $f(u_1, \dots, u_n) = v_2$ with $v_1 \neq v_2$, and its set of functions is not GR. Conversely, if $f(u_1, \dots, u_n) = v_1$ and $f(u_1, \dots, u_n) = v_2$ this set of functions is not GR and therefore there is no unique v with $f(u_1, \dots, u_n)$ GR. But if A is the statement 'a function is GR' and B is the statement 'there exists a set of functions which are GR', then

$$(A \Leftrightarrow B) \Leftrightarrow (\text{NOT } A \Rightarrow \text{NOT } B) \,\&\, (\text{NOT } B \Rightarrow \text{NOT } A)$$

is valid. Thus the theorem follows. □

Definition 14.4.6. The *characteristic function* χ_T of a set T is defined as $\chi_T(u) = 0$ when $u \notin T$ and $\chi_T(u) = 1$ when $u \in T$.

Definition 14.4.7. A set T is *recursive* if its characteristic function is GR.

Definition 14.4.8. If T is the empty set, or is the codomain of a GR function, then it is *recursively enumerable*.

14.4.9. [Quoted from 5Co66] "*There exists a recursively enumerable set which is not recursive*.

Proof. Our proof will specifically exhibit such a set. Let us first enumerate in any "effective" manner, all possible finite sets of equations in which a function symbol f of one variable occurs. Now enumerate all possible deductions from these equations. Since the nth system of equations defines a function, we shall denote this function by f_n. Even if f_n is not defined, we shall write $f_n(a) = b$ to mean that there is a deduction from the nth set of equations of the form $f(a) = b$, where a and b are natural numbers. We will now enumerate all deductions of the form $f_n(a) = b$. It can easily be shown that all our enumerations can be so chosen that there exist PR functions $\varphi_1, \varphi_2, \varphi_3$ such that if the equation $f_n(a) = b$ occurs in the mth place of our sequence, then $n = \varphi_1(m)$, $a = \varphi_2(m)$, $b = \varphi_3(m)$. Let g, h_1, h_2 be PR functions such that the map $(a, b) \to g(a, b)$ is a 1-1 correspondence between the set of all pairs of natural numbers and the set of natural numbers, and such that $a = h_1(g(a, b))$, $b = h_2(g(a, b))$. We now define F(n) as follows: Let $r = h_1(n)$ and $s = h_2(n)$. If there is a deduction of the equation $f_r(r + 1) = 0$ occurring in our list before the sth place, put $F(n) = r + 1$. If not, put $F(n) = 0$. Our first claim is that F is PR. This can be checked by explicitly writing out all the maps used. It should be fairly obvious that all the functions we have used were defined by quite simple recursion procedures. Observe that for all r, $r + 1$ occurs in the codomain if and only if $f_r(r + 1) = 0$ can be deduced. Let S be the codomain of F, χ its characteristic function. We claim that for no r can f_r be equal to χ. (If f_r is not defined this statement is vacuous). If f_r is defined then $f_r(r + 1)$ must be eventually computed. If $f_r(r + 1) = 0$ then $r + 1$ is in the codomain so $\chi(r + 1) = 1$. But if $f_r(r + 1) \neq 0$, then $r + 1$ is not in the range and $\chi(r + 1) = 0$. In either case $\chi(r + 1) \neq f_r(r + 1)$ and the theorem is proved". □

The reason this does not prove what is claimed is that the functions $f_n(a)$ are PR and therefore GR, the new distinct function F(n) together with its characteristic function χ defined in terms of the $f_n(a)$ is PR and GR, but because F(n) is a new function, we cannot guarantee that the set of functions $f_n(a)$, F(n) and χ together are GR, as was shown in the example previously given.

Because the conclusion of the proof is that we have a contradiction, we infer that the combination of $f_n(a)$, F(n) and χ together is not GR, even though they are separately.

Since the proof does not concern GR functions, it does not concern a recursively enumerable set S, since the codomain of S is not general recursive, nor is S the empty set. □

Thus we cannot conclude, but Gödel and Cohen do
14.4.10. "The decision procedure for Z_1 (equivalently X_1 or Z_2) is unsolvable". □

Theorem 14.4.11. *Every nonempty recursively enumerable set is in the codomain of* a PR *function.*

Proof. Let T be the codomain of a GR function f. If T is nonempty, now assume $t \in T$. Being GR, the function f is defined by a system of equations, so that there is a PR function g for which the if the mth deduction is in the form f(u) = v then g(m) = v, otherwise g(m) = t. The codomain of g is now T. □

14.5. Zermelo-Fraenkel set theory.

Sets are often used in the general description of mathematical systems; they include but are not restricted to collections of various types of number, for instance sequences or functions of numbers. The notion of a natural number itself can be expressed within this system. Further, we can introduce sets with an infinite number of members, or elements. We will not discuss here all aspects of generalisations even of sets. Such systems can for example be defined in category theory, and some are known as toposes.

This section gives a more detailed discussion of set theory than was provided in chapter III, section 2. A set of axioms for set theory was given for the first time by E. Zermelo in 1908, which was later developed by A. Fraenkel. This combined theory is often called Zermelo-Fraenkel (ZF) set theory and can be used to describe all of traditional mathematics. We will list its axioms again, and comment on them.

(1) *Axiom of extensionality*.
$\forall x, y\ (\forall z\ (z \in x \Leftrightarrow z \in y) \Rightarrow x = y)$.
As for the system Z_2, this states that a set is determined by its members.

We can now define the *includes* symbols.
$x \subseteq y \Leftrightarrow \forall z\ (z \in x \Rightarrow z \in y)$ means x is included in y, and can equal it.
$x \subset y \Leftrightarrow x \subseteq y$ & NOT $x = y$ means x is properly included in y; it does not equal y.

(2) *Axiom of the empty set*.
$\exists x\ \forall y$ (NOT $y \in x$).
This empty or null set is denoted by $\emptyset$.

(3) *Axiom of unordered pairs*.
$\forall x, y\ \exists z\ \forall w\ (w \in z \Leftrightarrow w = z$ OR $w = y)$.
The set z of unordered pairs is denoted by {x, y}. This means that {x} is {x, x}. For an ordered pair or *Cartesian product* <x, y>, we define this as the set {{x}, {x, y}}. It is possible to prove that <x, y> = <u, v> implies x = u and y = v.

A definition of a *function* can now be given.

A function is a set f of ordered pairs such that $\langle x, y\rangle$ and $\langle x, z\rangle$ in f $\Rightarrow$ y = z. The set of x is called the *domain* of the function and the set y the *codomain*, *image* or *target*. Some authors define the range to be the codomain of a function and others the domain, so we do not use this. A *map* or *mapping* is the set corresponding to the function itself. This mapping is *one-to-one*, an *injection* or a *monomorphism* (the last of such triples are used particularly for groups) if every $x' \in x$ maps to a unique $y' \in y$. A map is *onto*, a *surjection* or an *epimorphism* if every $y' \in y$ is the codomain of the function f with domain x. The mapping is *one-to-one and onto*, a *bijection* or an *isomorphism* if it is both injective and surjective.

(4) Axiom of the *union* of sets.

$\forall x\, \exists y\, \forall z\, (z \in y \Leftrightarrow \exists u\, (z \in u \,\&\, u \in x))$.

We call the set y the union of all the sets in x. Using axiom (3) we can then show that given x and y, there exists a z satisfying $z = x \cup y$, that is, $u \in z \Leftrightarrow u \in x$ OR $u \in y$.

(5) Axiom of *infinity*.

$\exists x\, (\emptyset \in x \,\&\, \forall y\, (y \in x \Rightarrow y \cup \{y\} \in x)$.

We have already come across this in Z_2: if x is a natural number, the successor of x is defined as $x \cup \{x\}$.

(6_n) Axiom of *restricted comprehension* (also called the axiom of *replacement*).

$\forall u_1, ..., u_k\, (\forall x\, \exists! y\, A_n(x, y, u_1, ..., u_k) \Rightarrow$

$\forall v\, \exists w\, \forall z\, (z \in w \Leftrightarrow \exists r\, (r \in v \,\&\, A_n(r, z, u_1, ..., u_k)))$.

This axiom scheme says that if for fixed $u_1, ..., u_k$ $A_n(x, y, u_1, ..., u_k)$ defines y uniquely as a function $\psi(x)$ of x, then for each v the codomain of v mapped to $\psi(v)$ creates a set. Whenever we define a set by "S is a set such that ..." we are using this axiom. Here the property A_n which defines ψ may be very non-constructive, for instance to verify it we may need to answer a question about infinite sets.

(7) Axiom of the *power set*.

$\forall x\, \exists y\, \forall z\, (z \in y \Leftrightarrow z \subseteq x)$.

This states that for every x there exists a set y of all subsets of x. This y is defined by a property which is not covered by the axiom of restricted comprehension, because it is not defined as the codomain of any function.

(8) Axiom of *choice*.

If $n \Rightarrow A_n \neq \emptyset$ is a function defined for all $n \in x$, then there exists another function f(n), and $f(n) \in A_n$.

This allows us to do an infinite amount of 'choosing' even without a property that would allow us to define the choice function which could otherwise be defined by axiom 6_n.

(9) Axiom of *regularity* (or *well-founding*).

$\forall x\, \exists y\, (x = \emptyset$ OR $(y \in x \,\&\, \forall z\, (z \in x \Rightarrow$ NOT $z \in y)))$.

This axiom is somewhat artificial and we include it for technical reasons only. It is never used in conventional mathematics. It states that every set $x \neq \emptyset$ contains a minimal element with respect to the $\in$ relation (but not $\subseteq$), because we want our sets built up from $\emptyset$ where all descending chains with respect to $\in$ terminate with $\emptyset$. This axiom is similar to the well-ordering property, except that $\in$ is not an ordering because we could have $x \neq y$, NOT $x \in y$ and NOT $y \in x$.

We will now give further comments on what happens if these axioms are dropped. If the axiom of extensionality does not hold then the system may contain atoms, which are sets x with $\forall y$ (NOT $y \in x$) but for which the sets x can be distinguished.

If the axiom of infinity does not hold, then we can take as a model of ZF the set of all finite sets which can be derived from $\emptyset$, which we have already mentioned for Z_2. Since the other axioms are also valid for finite sets, this axiom is independent of the other axioms of ZF, because axiom (5) becomes false for finite sets.

Axiom (6_n) for restricted comprehension enables the formation of a new set from all possible sets. This allows us to build sets up from smaller ones or cut sets down from larger ones. A replacement which only allows cutting sets down from larger ones is known as the axiom of separation:

($6'_n$) Axiom *of separation*.

$$\forall u_1, ..., u_k \; \forall x \; \exists y \; \forall z \; (z \in y \Leftrightarrow z \in x \; \& \; A_n(z, u_1, ..., u_k)).$$

Restating this, for every x there exists a set y comprising all z in x with property A_n. It can now happen that to verify A_n may be to ask a question about all sets, but this new set y will be a subset of x.

The axiom of separation results from that of restricted comprehension, which we restate as

$$\forall u_1, ..., u_k \; (\forall x \; \exists! y \; A'_n(x, y, u_1, ..., u_k) \Rightarrow$$
$$\forall v \; \exists w \; \forall z \; (z \in w \Leftrightarrow \exists r \; (r \in v \; \& \; A'_n(r, z, u_1, ..., u_k))),$$

which means we may set z = r. We can set y = x, because we have defined uniqueness as

$$\exists! y \; A'_n(x, y, u_1, ..., u_k) \Leftrightarrow \exists y \; \forall x((A'_n(x, x, u_1, ..., u_k) \Leftrightarrow y = x)),$$

and since x = x, the statement before the $\Rightarrow$ sign above is identically true. The separation axiom follows on putting $A'_n(x, x, u_1, ..., u_k) = A_n(x, u_1, ..., u_k)$. □

To prove that (6_n) is a stronger axiom than ($6'_n$), for every x denote P(x) as the power set of x, and put $t_0 = \emptyset$, $t_n = P(t_{n-1})$, with $M = \cup_n t_n$, so M is a model of all finite sets. If we put $D_0 = M$, $D_n = P(D_{n-1})$ and $N = \cup_n D_n$, then all the axioms of ZF hold in N with ($6'_n$) replacing (6_n), because when $x \in N$ also $P(x) \in N$. But we have just defined the function $n \rightarrow D_n$ in ZF, so if (6_n) was a valid axiom in N, then the infinite set $\{D_0, D_1, ...\}$ would be a member of the finite set N, which is a contradiction. □

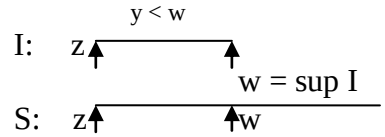
What we have to say on the relationship between the power set axiom and the axiom of infinity is dealt with in the section on ordinals and cardinals.

14.6. Well-ordering, ordinals and cardinals. [5Co66]

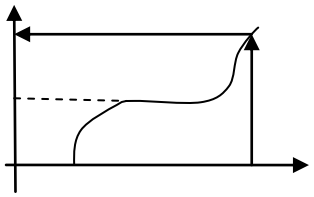
In this section we will first define a number of ideas we use to discuss the properties of well-ordering. A well-ordered set has a least element, but in proofs we will need to introduce the idea of an initial segment, which contains this least element and is less than a chosen element just outside of the segment.

(1) We define formally a *relation* R *on a set* X is a set of ordered pairs of members of X.
(2) $X \times Y$ is the set of all $<u, v>$ such that $u \in X$ and $v \in Y$. Note that if P(A) is the power set of A, then $X \times Y$ is a subset of $P(P(X \cup Y))$.
(3) $<x_1, ..., x_n> = <x_1, <x_2, ..., x_n>>$.

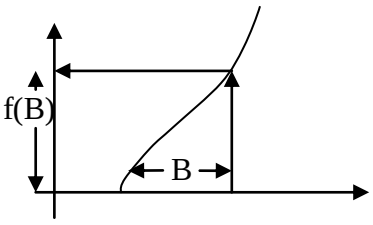
(4) If f is a function with domain x and $y \subseteq x$, then f | y, which means f *is restricted to* y, is the set of all <u, v> in f such that $u \in Y$.
(5) If $f(x) = f(y) \Rightarrow x = y$, then $f^{-1} = \{<x, y> \mid <y, x> \in f\}$.
(6) We will now write $x < y$ in place of <x, y>.
(7) We will write $x > y$ for $y < x$, and $x \leq y$ for $x < y$ OR $x = y$.
(8) A relation R on a set S *orders* S if
(i) $\forall x, y$ in S, one and only one of the following relations hold:
$x = y$ $\quad x < y$ $\quad y < x$
(ii) $x < y$ and $y < z \Rightarrow x < z$.
(9) A relation R *well-orders* S if R orders S and if
$I \subseteq S \;\&\; I \neq \emptyset \Rightarrow \exists z\,(z \in I \;\&\; \forall y\,(y \in I \Rightarrow \text{NOT } y < z))$.
This means R well-orders S if every subset of S has a least element with respect to R.
(10) If $I \subseteq S$, we write w = sup I, meaning w is the *supremum* of I, if w is the least member in S such that $y \in I \Rightarrow y < w$.



(11) If S is well-ordered by R, then I is an *initial segment* of S if $x \in I \;\&\; y < x \Rightarrow y \in I$.
This means I is an initial segment when I is all of S, or otherwise $\exists w$ with I the set of all y satisfying $y < w$. This is because if the complement of I in S is not $\emptyset$, suppose w is its least member. Then if $y < w$, y is in I, but if $y \in I$ and $w < y$ then by the property of an initial segment, $w \in I$, which contradicts the proposition that w is in the complement of I in S.
(12) A function f from an ordered set S to and ordered set T is *nondecreasing* if $x < y \Rightarrow f(x) \leq f(y)$ and *order preserving* or strictly monotonically increasing if $x < y \Rightarrow f(x) < f(y)$.



nondecreasing map



strictly monotonically increasing map or order preserving map

Theorem 14.6.1. *If* S *and* T *are well-ordered sets then at least one of* (i) *or* (ii) *holds*
(i) $\exists!$ f, *an order preserving surjective function, from* S *to an initial segment of* T.
(ii) $\exists$ *a unique order preserving surjective function from* T *to an initial segment of* S.

Proof. We will call an order preserving surjective map from a well-ordered set to an initial segment a *good* map.

A good map f satisfies for every x, $f(x) = \sup\{f(y): y < x\}$. This is because for the order preserving map, if $y < x$ then $f(y) < f(x)$, and so $f(x) \geq \sup\{f(y): y < x\} = t$, say. Now assume $f(x) > t$, with t in the codomain of f, is in an initial segment, and thus by the order preserving property contains $t = f(y)$ for some $y < x$, a contradiction, as was to be proved.

Two well-ordered sets S and T have at most identical maps between them, otherwise if there were two order preserving maps $f(x) \neq g(x)$ then we would obtain $\sup\{f(y): y < x\} = \sup\{g(y): y < x\}$, a contradiction.

The codomain of an order preserving map f from S to T satisfies $f(S) \subseteq T$.

Then for all $z > \sup\{x\}$

$$t_z = f(z) > f(\sup\{x\}) \in T,$$

and if there exists no f(w) with $f(\sup\{x\}) > f(w)$, then case (i) of the theorem is proved.

However, if there exists an $f(w) < f(\sup\{x\})$, since T is well-ordered, $\sup\{f(w)\}$ exists. Then the inverse map

$$f^{-1}(t_z) > f^{-1}(f(\sup\{x\})) = \sup\{x\} > f^{-1}\sup\{f(w)\}$$

has codomain $\subseteq S$, so $f^{-1}\sup\{f(w)\}$ becomes the supremum of S, and case (ii) of the theorem is proved. □

Theorem 14.6.2. *If* f *is a good injective map from* S *to* T *and* g *is a good injective map from* T *to* S, *then both* f *and* g *are bijective and therefore mutual inverses.*

Proof. Let f be the surjective map from S to the initial segment B of T. The restriction of g to B is also a good map f^{-1} from B to S, and since f^{-1} is surjective to S, since g is injective we have B = T and $f^{-1} = g$. □

Definition 14.6.3. For well-ordered sets S and T, we mean by $\widetilde{S} \leq \widetilde{T}$ that there is a good injective map from S to T. If this map is not surjective, we write $\widetilde{S} < \widetilde{T}$, and if it is bijective we write $\widetilde{S} = \widetilde{T}$.

We now introduce ordinal numbers. By the definition above they may be thought of as equivalence classes of well-ordered sets. The order relation defined there means that $\widetilde{S}$ and $\widetilde{T}$ are related by one and only one of the relations $\widetilde{S} < \widetilde{T}$, $\widetilde{T} < \widetilde{S}$ and $\widetilde{S} = \widetilde{T}$.

Here are some examples of well-ordered sets.

Example 14.6.4.

(i) The set $\mathbb{N}$ of natural numbers.

(ii) Let S and T be well-ordered sets with the intersection $S \cap T = \emptyset$. We can define S + T as $S \cup T$ under the ordering $w < x \Leftrightarrow (w \in S \,\&\, x \in T)$ OR $(w \in S \,\&\, x \in S)$ OR $(w \in T \,\&\, x \in T)$.

(iii) Let S and T be well-ordered sets. Define a well-ordering on the Cartesian product $S \times T$ by $\langle w, x\rangle < \langle y, z\rangle \Leftrightarrow x < z$ OR $(x = z \,\&\, w < y)$.

These definitions keep to the equivalence classes defined by =. For case (ii) we can define the sum of S and T not disjoint by putting them in the equivalence class with respectively S′ and T′ disjoint for which $\widetilde{S} = \widetilde{S'}$ and $\widetilde{T} = \widetilde{T'}$. □

Definition 14.6.5. A set x is called *transitive* when $z \in y \,\&\, y \in x \Rightarrow z \in x$.

Definition 14.6.6. An *ordinal* is a transitive set β that is well-ordered by ∈.

We can now select a standard representative from the equivalence class of well-orderings. We have already met the natural numbers as $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, etc. or equivalently defined by the relation $n + 1 = n \cup \{n\}$, and these sets n satisfy the condition of being ordinals. The reason we have introduced the transitive condition is so that this definition gives unique n. We remark that we can deduce from the axiom of regularity that a set ordered by $\in$ is well-ordered, but our discussion will be independent of this axiom.

Theorem 14.6.7. *An initial segment of an ordinal is an ordinal*.

Proof. If J is an initial segment of β, J is well-ordered by $\in$. Then if $y \in x$ and $x \in J$, then under the ordering relation $\in$ for β, $y < x$, which implies $y \in J$ and y is transitive. β is an ordinal because if $x \in \beta$, then x is the set of all y in β with x the successor of them. This follows from the $\in$ order relation and the fact that $x = \{y: y \in x\} = \{y: y \in x \,\&\, y \in \beta\}$. Now if β is an ordinal and $x \in \beta$, then x is an ordinal, since $x \subseteq \beta$, which implies x is well-ordered by $\in$. The transitive property implies that $z \in y \,\&\, y \in x \Rightarrow z \in x$, so that if $y \in \beta$ then $z \in \beta$. Also, if β is an ordinal then so is its successor, $\gamma = \beta \cup \{\beta\}$, and this satisfies the property that if x and y are in γ then x and $y \in \beta$, $x = \beta$ and $y \in \beta$, $x \in \beta$ and $y = \beta$, or finally $x = \beta$ and $y = \beta$, which implies $y \in x$, $x \in y$ or $x = y$. Thus γ is an ordinal. □

Theorem 14.6.8. *If* S *is a well-ordered set, there exists an ordinal* β *and a unique surjective order-preserving map* f *from* S *to* β, *so that* S *is bijective to an ordinal*.

Proof. We remark firstly that $\{\emptyset\}$ is the unique ordinal with one element, since $\emptyset$ is empty and it is the unique representative with no elements, so its successor which has one element $\{\emptyset\}$ is unique. Further, all successors are unique. Thus for any β, there is just one bijection of S to β. Let K be the set of all x in S with initial segment $I_u = \{y: y \leq u\}$ for every $u \leq x$, so that there is a bijection f_u to a unique ordinal $\beta(u)$. K is not empty, since if z is the least element of S, I_z is bijective to $\{\emptyset\}$ and $\{\emptyset\}$ is the unique ordinal with one element. K itself is an initial segment. Now consider the function f(x) with domain $x \in K$ given by $f(x) = f_x(x)$.

By the axiom of restricted comprehension, we allocate β as the codomain of f. Then β is an ordinal and f is a bijection from K to β. Now consider another bijection, g, from K to β. Then for $x \in K$, the restriction of g to I_x is a bijection from I*x* to an initial segment J of β where J is an ordinal, so by the definition of K $g(x) = f(x)$. This means f and β are unique. If K is all of S the proof is complete, so we assume $K \neq S$ and prove a contradiction. Let $y = \sup K$ and define g on $K \cup \{g\}$ by allocating for x in K $g(x) = f(x)$ and $g(y) = \beta$. This implies g maps I_y bijectively to the ordinal $\beta \cup \{\beta\}$, so we find the contradiction $y \in K$. □

Corollary 14.6.9. *If* β *and* γ *are ordinals, then* $\widetilde{\beta} = \tilde{\gamma}$ *implies* $\beta = \gamma$ *and if* $\widetilde{\beta} < \tilde{\gamma}$, $\beta < \gamma$.

Proof. The first part of the corollary is a restatement of the theorem. For the second part, if $\widetilde{\beta} < \tilde{\gamma}$ there is a bijection of β to a properly included initial segment of β, say $I = \{x < y\}$ for some $\gamma \in \beta$, but since I is an ordinal both $I = \beta$ and $I = \gamma$. □

In brief, we have demonstrated that *an ordinal is the set of all ordinals that precede it*.

Assertion 14.6.10. *If* β *is an ordinal, then* $\gamma = \beta \cup \{\beta\}$ *is the least ordinal greater than* β, *where* γ *by definition is* $\beta + 1$. □

Theorem 14.6.11. *If* S *is a set of ordinals, there exists a least* β *in* S.

Proof. Suppose $\gamma \in S$, then the β we want is the least member of the intersection $\gamma + 1 \cap S$. □

Theorem 14.6.12. *If* S *is a set of ordinals, there exists a least ordinal* δ *with* $\gamma \in S \Rightarrow \gamma < \delta$. *This* δ *is denoted by* sup S.

Proof. Suppose $\gamma \in S$ is its set of sums. If x, y and $z \in \gamma$, then for β_1, β_2 and $\beta_3 \in \gamma \in S$, $x \in \beta_1$, $y \in \beta_2$ and $z \in \beta_3$ are ordinals. For example, choose $\beta_1 < \beta_2 < \beta_3$, which gives x, y and $z \in \beta_3$, with $\in$ ordering γ. Let T be a subset of γ, with $\delta \cap T \neq \emptyset$ and $\delta \in S$. Since the least member of $\delta \cap T$ is also the least member in T, γ is well-ordered. It is transitive. This means γ is an ordinal for which $\delta \in S \Rightarrow \delta \in \gamma + 1$, where furthermore γ is the least of these ordinals. Thus $\delta = \gamma + 1 = \sup S$ is the required ordinal. □

Definition 14.6.13. γ is a *successor* if there exists a β with $\gamma = \beta + 1$. A *limit ordinal* satisfies $\gamma \neq 0$ and γ is not a successor.

Definition 14.6.14. β is a *natural number* if $\gamma \leq \beta \Rightarrow \beta$ is a successor.

Theorem 14.6.15. *A limit ordinal exists.*

Proof. This is a restatement of the axiom of infinity. Let x be a set satisfying this axiom, and $y = \sup\{x: x \text{ has domain all ordinals}\}$. Then y is a limit ordinal because $\delta \in x \Rightarrow \delta + 1 \in x$. □

For x a limit ordinal, we will put $\mathbb{N}$ as the least limit ordinal $\leq x$. For finite n, if n is a natural number then $x \in \mathbb{N}$. Conversely, if $y \in \mathbb{N}$ then $z \leq y$ implies z is either 0 or a successor, and y is a natural number. Thus $\mathbb{N}$ is the set of all natural numbers. □

Assertion 14.6.16. **(mathematical induction).** Let $x \subseteq \mathbb{N}$ and $\emptyset \in x$, then $n \in x \Rightarrow n + 1 \in x$ if and only if $x \in \mathbb{N}$. □

We will now extend mathematical induction to transfinite induction, where for ordinals γ the objective is to define a function $f(\gamma)$ when $f(\beta)$ has been obtained for every $\beta \leq \gamma$.

Theorem 14.6.17. **(transfinite induction)**. *Consider variables* u_i *and a function* A_n *such that* $\forall x\, \exists! y\, A_n(x, y, u_1, ..., u_k)$, *so that* A_n *defines a function* $y = e(x)$. *For every ordinal* γ *and a given set* z, $\exists! f$ *defined on the set* $\beta + 1$, *that is* $\{\beta: \beta \leq \gamma\}$, *satisfying* $f(0) = z$ *where for every* $\beta \leq \gamma$, $f(\beta) = e(h)$, *in which* h *is the function* f *restricted to* β.

Proof. Given γ and z, let S be the set of all ordinals $\beta \leq \gamma$ and f_β be the corresponding function. So if $\beta_1 \in S$ and $\beta_2 < \beta_1$ then $\beta_2 \in S$, where $f_{\beta 2}$ is $f_{\beta 1}$ restricted to $\beta_2 + 1$. This means S itself is an ordinal. If $\gamma \in S$ the theorem holds, otherwise we need to prove the following contradiction. Assume $\gamma \notin S$, and denote sup S by β_0. Now define the function g by $g(\beta) = f_\beta(\beta)$ for $\beta < \beta_0$. We now define $f(\beta) = g(\beta)$ for $\beta < \beta_0$ with $f(\beta_0) = e(g)$, so that this unique f satisfies $\beta_0 \in S$, in violation of what we assumed. □

Example 14.6.18. We give examples of the function e(x).

(i) Define $\gamma + \delta$ for fixed γ by setting $\gamma + 0 = \gamma$ and $\gamma + \delta$ by $\sup\{\gamma + \beta: \gamma < \delta\}$.
(ii) Define $\gamma.\delta$ for fixed γ by $\gamma.0 = 0$, $\gamma.(\delta + 1) = \gamma.\delta + \delta$ and $\gamma.\delta$ by $\sup\{\gamma.\beta: \gamma < \delta\}$.
(iii) Define $\gamma\,^n|\delta$ for fixed γ by $\gamma\,^n|0 = 1$, $\gamma\,^n|(\delta + 1) = (\gamma\,^n|\delta)\,^{n-1}|(\gamma\,^n|w_n)$ and $(\gamma\,^n|\delta) = \sup\{\gamma\,^n|\delta: \gamma < \delta\}$.
(iv) Define $\gamma|^n\,\delta$ for fixed γ by $\gamma|^n\,0 = 1$, $\gamma|^n\,(\delta + 1) = (\gamma|^n\,\delta)|^{n-2}\,(\hat{\gamma}|^n\,(\delta + 1))$ and $(\gamma|^n\,\delta) = \sup\{\gamma|^n\,\delta: \gamma < \delta\}$.

If γ and δ are ordinals, so are $\gamma + \delta$, $\gamma.\delta$, $\gamma \uparrow \delta$, $\gamma\,^n|\delta$ and $\gamma\,|^n\,\delta$.

Definition 14.6.19. An ordinal is called *countable* if there exists at least one bijective map between itself and the natural numbers, $\mathbb{N}$.

If γ is countable, then because $\delta < \gamma \Leftrightarrow \delta \in \gamma$, we can enumerate all ordinals less than γ. We will show that if γ and δ are countable, so are $\gamma + \delta$, $\gamma.\delta$, $\gamma\uparrow\delta$, $\gamma\,^{n}|\delta$ and $\gamma|^{n}\,\delta$. Indeed, the cases given of $\gamma\,^{n}|\delta$ and $\gamma|^{n}\,\delta$ subsume the others. If δ is the least countable ordinal such that $\gamma\,^{n}|\delta$ is not countable, then by relation (k) of 16.3, if $\delta = \beta + 1$, $\gamma\,^{n}|\delta = \gamma\,^{n}|(\gamma\,^{n-1}|\beta)$ is countable, whilst if δ is a limit ordinal and γ_i are all its predecessors $\gamma\,^{n}|\delta = \sup\{\gamma\,^{n}|\gamma_i\}$ and hence is the union of countably many countable sets and is countable. The situation for $\gamma|^{n}\,\delta$ follows from the relation (l) of 14.3, namely $\gamma|^{n}\,\delta = (\gamma|^{n}\,\delta)|^{n-2}\,(\hat{\gamma}|^{n}\,\delta)$. □

We now discuss cardinal numbers. Cardinals are ordinals and are equal if there is a bijection between their members. This implies if $\overline{\overline{b}}$ and $\overline{\overline{d}}$ are countable cardinals there exists an arithmetic in which $\overline{\overline{b}} + \overline{\overline{d}}$, $\overline{\overline{b}}.\,\overline{\overline{d}}$, $\overline{\overline{b}}\uparrow\overline{\overline{d}}$ etc., are also.

Definition 14.6.20. If A and B have the same *cardinality*, written $\overline{\overline{A}}$ and $\overline{\overline{B}}$ respectively, then there is a bijective mapping from A to B, and we write $\overline{\overline{A}} = \overline{\overline{B}}$.

Thus $\overline{\overline{A}} = \overline{\overline{B}}$ defines an equivalence relation. □

Definition 14.6.21. $\overline{\overline{A}} \leq \overline{\overline{B}}$ defines an injective map from A to B.

Theorem 14.6.22. (Schröder-Bernstein). $\overline{\overline{A}} \leq \overline{\overline{B}}$ *and* $\overline{\overline{B}} \leq \overline{\overline{A}}$ *implies* $\overline{\overline{A}} = \overline{\overline{B}}$.

Proof. This follows from the properties of injective and bijective maps. The explicit proof in Cohen does not use the axiom of choice. □

Definition 14.6.23. If $\overline{\overline{A}} \leq \overline{\overline{B}}$ and NOT $\overline{\overline{A}} = \overline{\overline{B}}$, we say $\overline{\overline{A}} < \overline{\overline{B}}$.

Thus $\overline{\overline{A}} < \overline{\overline{B}}$ is a partial ordering of A and B, in the sense of chapter III section 5. □

Theorem 14.6.24. (well-ordering). *The axiom of choice holds if and only if every set* S *can be well-ordered*.

Proof. Assuming the axiom of choice on the identity function of the power set P(S), it follows that there exists a function f for the set of all $x \subseteq S$ with $x \neq \emptyset$ such that $f(x) \in x$. Denote a good well ordering of a set $T \subseteq S$ a well-ordering if for every $x \in T$, x is this function on the complement of y in S, where $\{y: y \in T \;\&\; y < x\}$. This means we have chosen a 'next' element in the well-ordering by applying the choice function to those elements remaining, so that by successively choosing elements, we are able to well-order all of S.

In the reverse direction, a well-ordering implies the axiom of choice because if we well-order a set $S = \bigcup_z A_z$, on defining f(z) as the least element in A_z we get a choice function. □

Definition 14.6.25. A *cardinal* is the least ordinal d, where for every ordinal b, $b < d$ implies $\overline{\overline{b}} < \overline{\overline{d}}$.

This means we have chosen a cardinal d as a representative from the equivalence class $\overline{\overline{d}}$ of sets of the same cardinality, where if we apply the axiom of choice we can well-order every set of ordinals.

We often write 2^A for the cardinality of the power set P(A) of A, by analogy with finite sets. If the ordinals A and B are countable, we have already seen that B^A is countable, and therefore so is the cardinal 2^A. □

It is instructive to compare what we are saying on the uncountable continuum hypothesis against other approaches. We will choose amongst other texts P. J. Cohen's *Set theory and the continuum hypothesis* [5Co66], page 67. Let $\overline{\overline{P(A)}}$ be the cardinal of the power set of A. Cohen gives Cantor's theorem, and it is an accurate reflection of Cantor's argument, as

14.6.26. "*For any* A, $\overline{\overline{A}} < \overline{\overline{P(A)}}$",

with the following argument:

"*Proof*. Clearly $\overline{\overline{A}} \leq \overline{\overline{P(A)}}$. Assume θ maps A onto P(A). Let z = {x: NOT x ∈ θ(x)}. Then z ⊆ A and if z = θ(y) for y ∈ A then y ∈ z ⇒ NOT y ∈ θ(y) ⇒ NOT y ∈ z, and NOT y ∈ x ⇒ y ∈ θ(y) ⇒ y ∈ z, which is clearly impossible." □

We comment that if $\overline{\overline{A}} = \overline{\overline{P(A)}}$ then if it is intended that NOT x is derived from the set of all x, then z = ∅. Then if y ∈ z, y = ∅, and if z = θ(∅), which it does not, since A ∉ ∅, then we confirm its impossibility, so we still have $\overline{\overline{A}} = \overline{\overline{P(A)}}$.

On the other hand, if θ(x) is a member obtained by a choice of NOT x, then the argument is independent of whether or not we are mapping to the power set P(A), that is, it should work for any such function with codomain θ(x). It clearly does not, since such functions can be consistent.

Thus the proof does not demonstrate that $\overline{\overline{A}} \neq \overline{\overline{P(A)}}$. □

For the theory we are developing, where for countably infinite sets $\overline{\overline{A}} = \overline{\overline{P(A)}}$, it follows that we are not making any essential distinction between countable ordinals and their cardinals. Indeed if β and δ are countable ordinals, then they are bijective for some map with ℕ, ℕ is bijective with $\beta\,{}^{n}\Uparrow\delta$ and $\beta\Uparrow^{n}\,\delta$, their corresponding cardinals are defined by a bijection, and for a countable cardinal, this means it is bijective with ℕ also.

The cardinals have an arithmetic in which for finite disjoint $\overline{\overline{A}}$ and $\overline{\overline{B}}$, $\overline{\overline{A}} + \overline{\overline{B}} = \overline{\overline{A}} \cup \overline{\overline{B}} = \overline{\overline{A}} \,.\, \overline{\overline{B}} = \overline{\overline{A}} \times \overline{\overline{B}} = \max(\overline{\overline{A}}, \overline{\overline{B}})$. This assertion can be proved using the axiom of choice. □

14.7. Gödel's completeness theorem. [5Co66]

Consider a collection S of statements using constants c_a and relation symbols R_b, where each R_b has a fixed number of variables. For a non-empty set M, let the mapping from constant symbols to elements of M be $c_a \to c'_a$, and similarly $R_b \to R'_b$ be the k-fold map of relation symbols between the M × M × ..., × M. To every statement using these constant and relation symbols in M, we will speak of its 'truth value' under this interpretation.

Definition 14.7.1. For a formula A, assumed a good wff with free variables $\{u_1, \ldots u_n\}$ and with $\{u'_1, \ldots u'_n\} \in M$, we define the truth value of A in M at $\{u'_1, \ldots u'_n\}$ by

(1) If A is of the form $u_i = u_j$, $u_i = c$ or $c_i = c_j$ respectively, then A is true at $u_1', \ldots u'_n$ if $u'_i = u'_j$, $u'_i = c'$ or $c'_i = c'_j$.

(2) When A is the m-ary relation symbol $R(t_1, \ldots t_m)$, where each t_k is a constant symbol or one of $u_1, \ldots u_n$, then A is true at $u_1', \ldots u'_n$ if the m-tuple $(t_1, \ldots t_m)$ is given by the mapping from R to R′ under the interpretation.

(3) For a function of formulas A in the propositional calculus, evaluate A at $u_1', \ldots u'_n$ by the rules of the propositional calculus.
(4) When A is of the form $\exists v\, B(v, u_1, \ldots u_n)$ then A is true at $u_1', \ldots u'_n$ if for some v' in M, $B(v, u_1, \ldots u_n)$ is true at $v', u'_1, \ldots u'_n$.
(5) When A is of the form $\forall v\, B(v, u_1, \ldots u_n)$ then A is true at $u_1', \ldots u'_n$ if for every v' in M, $B(v, u_1, \ldots u_n)$ is true at $v', u'_1, \ldots u'_n$.

We note that when the number of variables, n, is zero, this just defines truth in M.

Definition 14.7.2. For a collection S of statements using constants c_a and relation symbols R_b, a *model* for S is a set M with an interpretation of some, possibly the empty set, of constant and relation symbols satisfying $c_a \rightarrow c'_a$, and $R_b \rightarrow R'_b$ above.

Definition 14.7.3. A set of statements S is *consistent* if and only if the statement A & NOT A cannot be derived from S for any A.

Assertion 14.7.4. *If* A *is a valid statement, it is true in every model, and if a set of statements* S *has a model then it is consistent*. □

Lemma 14.7.5. *Given an arbitrary statement* A *and a consistent set of statements* T, *then either* T ∪ {A} *or* T ∪ {NOT A} *is consistent*.

Proof. If T ∪ {A} is inconsistent, then for some B_k in T and some C, A & B_1 & ... & $B_n \Rightarrow$ C & NOT C is valid. If also T ∪ {NOT A} is inconsistent, then for some B'_k in T and some C, (NOT A) & B'_1 & ... & $B'_{n'} \Rightarrow$ C & NOT C is valid. Using the propositional calculus, we find that B_1 & ... & B_n & B'_1 & ... & $B'_{n'} \Rightarrow$ C & NOT C is valid, so that T is inconsistent. □

Theorem 14.7.6. **(completeness of the propositional calculus)**. *If a set of statements* S *with no quantifiers is consistent, then there exists a model* M *for* S *where every element of* M *is of the form* c'_a *for some constant symbols* c_a *occurring in* S.

Proof. Let S be a well-ordered set of statements, which entails that there is a well-ordering of all constant and relation symbols appearing in S. This then implies for constant and relation symbols c_i and R_b respectively that there is a well ordering of all statements of type $c_i = c_j$ and $R_b(c_1, \ldots c_n)$ in S. Denote these statements by V_a, and use the following induction procedure on a to define statements W_a. If V_a is consistent with $\{W_b: b < a\}$, select $W_a = V_a$, and if not, put W_a = NOT V_a. By lemma 14.8.5 and induction on a, for every a S ∪ $\{W_b: b \leq a\}$ is consistent. Because any contradiction must be obtained from a finite number of statements, we deduce that U = S ∪ $\{W_a\}$ is a consistent system. For every c_i in S, $c'_i = c_j$ when j is the least index with $c_i = c_j$ in U. Such a constant exists because there are c_i satisfying $c_i = c_i$. If M is the set of c'_i, define R'_j as the set of all $<c'_{i1}, \ldots c'_{in}>$ with $R_j(c'_{i1}, \ldots c'_{in})$ in U, and thus the model M is a subset of c_i. A look at rule C for equality in section 2 shows that any ambiguity in determining R'_j results in a contradiction in U. We now see that every statement in U, or its negation, is a consequence of the W_a, since any relation in S or its negation appears in W_a. The negation cannot occur because U would then be inconsistent. Since we have defined our model M so that all W_a are true, it follows that the statements of U and thus S are also true. □

Let T be a system of statements, each of which has either no quantifiers or starts with a quantifier. From rules D and G of section 2, we see this is not a significant restriction. To form a new system, if the statement $\forall x\, A(x)$ and the constant symbol c are present in T, then

include the statement A(c). For ∃x A(x), choose a c outside of T, and adjoin the statement A(c). Call the resulting system T*, so that T is now a subset of T*.

Lemma 14.7.7. *If* T *is consistent, so is* T*.

Proof. By the application of rule E of section 2, this is a restatement of the consistency of T* under the result that

$$\forall x\ A(x) \Rightarrow A(c)$$

is valid. By the use of rule F of section 2, A(c) ⇒ B is valid implies

$$\exists x\ A(x) \Rightarrow B,$$

so if a contradiction can be reached with an adjoined statement, it can be reached without it. Rule F has been designed to reach this conclusion. □

Theorem 14.7.8. **(Gödel completeness)**. *Let* S *be any consistent set of statements. Then there exists a model* M *for* S *whose cardinality does not exceed the cardinality of the number of statements in* S *when* S *is infinite, and whose cardinality is finite when* S *is finite*.

Proof. Rules D and G of section 2 allow the replacement of an arbitrary system T by a defined equivalent system T*. From the consistent system S, we define S_n for each natural number n by putting $S_0 = S$ and $S_{n+1} = S^*_n$. Define $S' = \cup_n S_n$ so that M is the model of the theorem for that subset of S′ without quantifiers. Now suppose A is a statement in S with q quantifiers, and assume that any statement in S′ with fewer than q quantifiers is true in M. The theorem holds if q = 0. If this is not the case, let us firstly consider the case when A is of the form ∀x D(x). Then any element of M is of the form c'_i for the corresponding constant symbol c_i in S. Since the S_n are increasing, both A and c_i are in some S_k, so that $D(c_i)$ occurs in S^*_k, and thus in S′, with fewer than q quantifiers. This means for any c'_i in M, $D(c_i)$ is true in M, and so is ∀x D(x). Next, when A is of the form ∃x D(x), we can deal with this case analogously. Since the cardinality of any system T* derived from T is the same as T when T is infinite, and is finite when T is finite, the theorem follows. □

14.8. A remark on Gödel incompleteness.

The Formalist programme of showing the consistency of various systems was affected by Gödel's incompleteness theorem, which states that the consistency of a mathematical system cannot be proved except by methods more powerful than those of the system itself.

14.8.1. "**Gödel's incompleteness theorem**". *The consistency of* Z_1 *cannot be proved in* Z_1.

The second proof given by Gödel of the incompleteness theorem relies in a direct way on the diagonal argument for PR functions, which we have deconstructed in section 4. In particular, if we enumerate all PR formulas, as happens in this proof, then for example we enumerate f(x) = 0 and f(x) = 1, so the result is inconsistent, which is not allowed in Z_1 ($0 \neq 1$). If one type of PR function is chosen, this may eliminate the choice of other PR functions. This does not indicate Z_1 is inconsistent or contains unprovable statements, since Z_1 disallows the eliminated choices. □

14.9. Exercises.

Prove the step-down equation for right nest operators of section 14.3

$$\forall x, y, n\ x\ |^n\ (y + 1) = (x\ |^n\ y)\ |^{n-2}\ (\hat{x}\ |^n\ (y + 1)).$$