

Part A

Elementary Theory

Chapter I

Preliminaries

The fundamental set theory of this book is Zermelo-Fraenkel set theory. In this chapter we give a brief account of this theory, insofar as we need it. Sections 1 through 5 cover the early development of the theory up to ordinal and cardinal numbers. The remaining six sections deal with some special topics of direct relevance to the subject matter of this book, and the coverage is therefore a little more complete than in the previous sections.

1. The Language of Set Theory

The language of set theory, LST, is the first-order language with predicates = (equality) and \in (set membership), logical symbols \wedge (and), \neg (not), and \exists (there exists), variables v_0, v_1, \dots , and (for convenience) brackets $(,)$.

The *primitive* (or *atomic*) formulas of LST are strings of the forms

$$(v_m = v_n), \quad (v_m \in v_n).$$

The *formulas* of LST are generated from the primitive formulas by means of the following schemas: if Φ, Ψ are formulas, so too are the strings

$$(\Phi \wedge \Psi), \quad (\neg \Phi), \quad (\exists v_n \Phi).$$

(We generally use capital Greek letters to denote formulas of LST.)

The notions of *free* and *bound* variables are defined as usual. A *sentence* is a formula with no free variables.

We write $x \notin y$ for $\neg (x \in y)$ and $x \neq y$ for $\neg (x = y)$. (We generally use, x, y, z , etc. to denote arbitrary variables of LST.)

The defined logical symbols $\vee, \rightarrow, \leftrightarrow, \forall$ are introduced in the usual way, and are frequently treated as if they were basic symbols of LST (i.e. having the same status as \wedge, \neg, \exists). Likewise for the bounded quantifiers $(\exists v_m \in v_n)$ and $(\forall v_m \in v_n)$ (where $m \neq n$), introduced by the schemas:

$$(\exists v_m \in v_n) \Phi \quad \text{replaces} \quad \exists v_m ((v_m \in v_n) \wedge \Phi);$$

$$(\forall v_m \in v_n) \Phi \quad \text{replaces} \quad \forall v_m ((v_m \in v_n) \rightarrow \Phi).$$

The symbols \subseteq and $\exists!$ are defined thus:

$$\begin{aligned} y \subseteq z & \quad \text{abbreviates } (\forall x \in y)(x \in z); \\ \exists! x \Phi & \quad \text{abbreviates } \exists y \forall x (y = x \leftrightarrow \Phi). \end{aligned}$$

(Thus $\exists! x \Phi$ means “there is a unique x such that Φ ”.) We also write

$$y \subset z \quad \text{to mean } y \subseteq z \wedge y \neq z.$$

The above abbreviations are never regarded as a fundamental part of the language LST, however, unlike the bounded quantifiers, etc.

One final remark. In writing formulas, we strive for legibility at the expense of strict adherence to the syntax of LST. This particularly applies to our use of parentheses, which are omitted wherever possible. Also, when nesting of clauses is required, we sometimes use both (square) brackets as well as parentheses, for clarity. Out notation for the interpretation of variables in formulas is also chosen with clarity in mind. If we write, say, $\Phi(v_i, v_j)$, we mean that the free variables of Φ are amongst the variables v_1, v_j . If we subsequently write $\Phi(x, y)$, where x and y are specific sets, we mean that Φ is a valid assertion when x interprets v_i and y interprets v_j . (Of course, we have also decided to use x, y, z , etc. to denote arbitrary variables of LST. But in any given case, the context should indicate the intended meaning.¹)

2. The Zermelo-Fraenkel Axioms

The theory ZF is the LST theory whose axioms are the usual axioms for first-order logic (for the language LST), together with the following axioms (i)–(vii):

- (i) Extensionality: $\forall x \forall y [\forall z (z \in x \leftrightarrow z \in y) \rightarrow (x = y)]$
- (ii) Union: $\forall x \exists y \forall z [z \in y \leftrightarrow (\exists u \in x)(z \in u)]$
- (iii) Infinity: $\exists x [\exists y (y \in x) \wedge (\forall y \in x)(\exists z \in x)(y \in z)]$
- (iv) Power Set: $\forall x \exists y \forall z [z \in y \leftrightarrow z \subseteq x]$
- (v) Foundation: $\forall x [\exists y (y \in x) \rightarrow \exists y (y \in x \wedge (\forall z \in y)(z \notin x))]$
- (vi) Comprehension (schema): $\forall \vec{a} \forall x \exists y \forall z [z \in y \leftrightarrow z \in x \wedge \Phi(z, \vec{a})]$,

where Φ is any LST formula whose free variables are amongst z, \vec{a} , and where the variables \vec{a}, x, y, z are all distinct.

(We use \vec{x}, \vec{a} , etc. to denote finite strings of variables, $\forall \vec{a}$ to abbreviate $\forall a_1, \dots, \forall a_n$ and $\Phi(z, \vec{a})$ to abbreviate $\Phi(z, a_1, \dots, a_n)$. In more complicated situations,

¹ Strictly speaking there is no clash of notation here. As far as formal set theory is concerned there are simply *variables* (to denote “sets”). But as usual, to avoid incomprehensible use of quantifiers and formulas to define specific sets, we argue in a loose, *semantic* fashion whenever possible, and then it can be useful to distinguish between “formal variables” and “sets which interpret those variables”.

we often use expressions such as $\vec{x}_0, \dots, \vec{x}_n$. Here, \vec{x}_0 will denote some sequence x_{00}, \dots, x_{0k} , \vec{x}_1 will be another sequence x_{10}, \dots, x_{1l} , possibly of a different length, according to context, and so on.)

(vii) Collection (schema):

$$\forall \vec{a} [\forall x \exists y \Phi(y, x, \vec{a}) \rightarrow \forall u \exists v (\forall x \in u) (\exists y \in v) \Phi(y, x, \vec{a})],$$

where Φ is any LST-formula whose free variables are amongst y, x, \vec{a} , and where the variables \vec{a}, x, y, u, v are all distinct.

In (iii), the exact formulation of the Axiom of Infinity is not important, and different texts often give different formulations. The main point is to guarantee the existence of at least one infinite set. Axiom (vi) (the Comprehension Axiom schema) is sometimes referred to as the Subset Selection schema. The German word *Aussonderungssaxiom* is also quite common for this axiom scheme. In Axiom (vii) (Collection), notice that we have placed the variable y before the variable x . This is purely a stylistic convention, of course, and reflects the fact that in our representation of a function as a set of ordered pairs, we shall take the first member of each ordered pair as the value of the function and the second element as the argument. Axiom schemas (vi) and (vii) are often replaced by a single schema: the *Axiom of Replacement*.

Notice that by virtue of the two axiom schemas, the above list of axioms for ZF is infinite. We shall soon be able to prove that no finite collection of LST sentences suffices to axiomatise ZF.

By the Axiom of Infinity, there exists at least one set. The Axiom of Comprehension then yields the existence of the empty set \emptyset . Many texts include as an axiom of ZF the *Null Set Axiom*, which is the assertion that there exists a set having no elements, viz.:

$$\exists x \forall y (y \notin x).$$

Zermelo-Fraenkel set theory includes one further axiom:

(viii) Axiom of Choice (AC):

$$\begin{aligned} \forall x [(\forall y \in x) (y \neq \emptyset) \wedge (\forall y, y' \in x) (y \neq y' \rightarrow \forall w (w \in y \leftrightarrow w \notin y')) \\ \rightarrow (\exists z) (\forall y \in x) (\exists! v \in y) (v \in z)]. \end{aligned}$$

We denote Zermelo-Fraenkel set theory (which includes AC) by ZFC. This nomenclature is now fairly standard, despite the rather unfortunate fact that it means that the letters ZF do not stand for “Zermelo-Fraenkel” set theory, but just a part of that theory. To try to avoid any confusion, throughout the book we shall stick to the abbreviated notations ZF and ZFC. Hence, we shall have the “equation”

$$\text{ZFC} = \text{ZF} + \text{AC}.$$

ZFC is our basic set theory. On occasions it will be important to note that AC is not being used in an argument, and in such cases we shall write, for example,

$$\text{ZF} \vdash \Phi$$

or else

$$\Phi \rightarrow_{\text{ZF}} \Psi$$

to mean, respectively, that Φ is provable in ZF or that Ψ is provable from Φ together with the axioms of ZF.

3. *Elementary Theory of ZFC*

3.1 (Sets and Classes). The basic objects of discussion of ZFC (i. e. the objects over which the variables range) are called *sets*. The *universe* is the collection of all sets, and is denoted by V . If $\Phi(v_0, v_1, \dots, v_n)$ is an LST formula and x_1, \dots, x_n are sets, the collection of all sets x for which $\Phi(x, x_1, \dots, x_n)$ is a *class*, denoted by

$$\{x \mid \Phi(x, x_1, \dots, x_n)\}.$$

Every set, y , is a class (consider the formula $\Phi(x, y) \equiv (x \in y)$), but not every class is a set (consider the formula $\Phi(x) \equiv (x \notin x)$, which would lead at once to the Russell paradox if the class it defined were a set). We often write

$$\{x \in y \mid \Phi(x, x_1, \dots, x_n)\}$$

in place of

$$\{x \mid x \in y \wedge \Phi(x, x_1, \dots, x_n)\}.$$

(By the Axiom of Comprehension, this class is always a set.) We generally use capital Roman letters X, Y, Z etc. to denote classes, with lower case Roman letters being reserved for sets (as well as for variables of LST, which denote sets, of course). A class which is not a set is called a *proper class*. Proper classes do not fall under the scope of the axioms of ZFC, but their usage is convenient. We assume the reader is familiar both with the use of proper classes in set theory and the means by which such usage may be avoided if required. A particular example occurs in VI.1, where we discuss the rudimentary functions. It is convenient, though avoidable, to develop the relevant theory in terms of “functions” defined on the whole of V , even though, as proper classes these cannot be *functions* in the sense of set theory at all.

Our set-theoretic notation is standard. The set consisting of precisely the elements x_1, \dots, x_n is denoted by

$$\{x_1, \dots, x_n\};$$

$\{x\}$ is the *singleton* of x , and $\{x, y\}$ is the *unordered pair* of x, y . Many texts include as an axiom of ZF the *Pairing Axiom*, which asserts that for every pair of elements x, y , the set $\{x, y\}$ exists, i. e.

$$\forall x \forall y \exists z \forall u (u \in z \leftrightarrow u = x \vee u = y).$$

However, as this “axiom” is easily proved from the axioms we listed earlier, we did not take it as a basic axiom.

The *ordered pair* of x and y is defined by

$$(x, y) = \{\{x\}, \{x, y\}\},$$

and has the property that

$$(x, y) = (x', y') \quad \text{iff } x = x' \text{ and } y = y'.$$

The *union* of x (i.e. the set of all members of all members of x) is denoted by $\bigcup x$, and is guaranteed to exist by the Union Axiom. We write $x \cup y$ instead of $\bigcup \{x, y\}$. The *intersection* of x , $\bigcap x$, is defined by

$$y \in \bigcap x \quad \text{iff } (\forall z \in x)(y \in z),$$

and is a set whenever $x \neq \emptyset$. (By our definition, $\bigcap \emptyset = V$, but this is not a case that will ever concern us.) We write $x \cap y$ for $\bigcap \{x, y\}$. The *difference* of x and y is defined by

$$x - y = \{z \in x \mid z \notin y\}.$$

The *power set* of x (i.e. the set of all subsets of x) is denoted by $\mathcal{P}(x)$, and is guaranteed to exist by the Power Set Axiom.

3.2 (Ordinals). A class M is said to be *transitive* if

$$x \in y \in M \rightarrow x \in M.$$

If $\text{Trans}(v_0)$ denotes the LST formula

$$(\forall v_1 \in v_0)(\forall v_2 \in v_1)(v_2 \in v_0),$$

then a set x will be transitive iff $\text{Trans}(x)$.

An *ordinal number* (or simply, an *ordinal*) is a transitive set which is linearly ordered by \in . We use $\alpha, \beta, \gamma, \dots$ to denote ordinals. We denote by $\text{On}(v_0)$ the LST-formula

$$\text{Trans}(v_0) \wedge (\forall v_1 \in v_0)(\forall v_2 \in v_0)(v_1 = v_2 \vee v_1 \in v_2 \vee v_2 \in v_1).$$

It is not hard to show that a set x will be an ordinal iff $\text{On}(x)$.

If α, β are ordinals, either $\alpha = \beta$ or $\alpha \in \beta$ or $\beta \in \alpha$. So the class

$$\text{On} = \{x \mid \text{On}(x)\}$$

is totally ordered by \in . We often write $\alpha < \beta$ instead of $\alpha \in \beta$, and $\alpha \leq \beta$ instead of $(\alpha < \beta \vee \alpha = \beta)$. It is easily seen that $\alpha < \beta$ is equivalent to $\alpha \subset \beta$. Moreover, for any ordinal α ,

$$\alpha = \{\beta \mid \beta < \alpha\}.$$

By the Axiom of Foundation, the relation $<$ is in fact a well-ordering of On (i.e. every non-empty subset of On has a $<$ -least element).

If A is a set of ordinals, then $\cup A$ is also an ordinal. In fact, $\cup A$ is the least ordinal δ such that $(\forall \alpha \in A)(\alpha \leq \delta)$. This least δ is also called the *supremum* of A , denoted by $\text{sup}(A)$. Thus $\text{sup}(A)$ and $\cup A$ coincide.

The first ordinal (under the canonical well-ordering \in) is the null set, \emptyset , but when considered as an ordinal it is usually denoted by 0 . The next ordinal is the set $\{0\}$, denoted by 1 . Then comes the ordinal $\{0, 1\}$, denoted by 2 , followed by $3 = \{0, 1, 2\}$, and so on. If α is an ordinal, so too is $\alpha \cup \{\alpha\}$, and there is no ordinal γ strictly between α and $\alpha \cup \{\alpha\}$. We call $\alpha \cup \{\alpha\}$ the *successor* of α , denoted by $\alpha + 1$. Any ordinal of the form $\alpha + 1$ is called a *successor ordinal*. An ordinal α is a successor ordinal iff $\text{succ}(\alpha)$, where $\text{succ}(v_0)$ is the LST-formula

$$\text{On}(v_0) \wedge (\exists v_1 \in v_0)(\forall v_2 \in v_0)(v_2 \in v_1 \vee v_2 = v_1).$$

A non-zero ordinal which is not a successor ordinal is called a *limit ordinal*. If $\text{lim}(v_0)$ is the LST-formula

$$\text{On}(v_0) \wedge (\exists v_1 \in v_0)(v_1 = v_1) \wedge (\forall v_1 \in v_0)(\exists v_2 \in v_0)(v_1 \in v_2),$$

then an ordinal α will be a limit ordinal iff $\text{lim}(\alpha)$. Using the Axiom of Infinity, together with other ZF axioms, it can be shown that a limit ordinal exists. The least limit ordinal is denoted by ω . The elements of the set ω are precisely the finite ordinal numbers, and are called the *natural numbers*. We usually denote natural numbers by m, n, i, j, k , etc. Notice that ω is definable by the formula

$$\text{lim}(v_0) \wedge (\forall v_1 \in v_0)(\text{succ}(v_1) \vee (\forall v_2 \in v_1)(v_2 \neq v_2)).$$

We usually write $\exists \alpha \Phi(\alpha)$ in place of

$$\exists v_0 [\text{On}(v_0) \wedge \Phi(v_0)],$$

and $\forall \alpha \Phi(\alpha)$ in place of

$$\forall v_0 [\text{On}(v_0) \rightarrow \Phi(v_0)].$$

If $(X, <)$ is a well-ordered set, there is a unique ordinal number α such that $(X, <)$ is isomorphic to α (with the usual ordering). This α is called the *order-type* of $(X, <)$, denoted by $\text{otp}(X, <)$.

3.3 (Relations and Functions). Let $n > 0$ be a natural number. The n -tuple (x_1, \dots, x_n) of sets x_1, \dots, x_n is defined thus:

$$\begin{aligned} \text{if } n = 1, & \quad (x_1) = x_1; \\ \text{if } n > 1, & \quad (x_1, \dots, x_n) = (x_1, (x_2, \dots, x_n)) \\ & \quad (= (x_1, (x_2, (x_3, \dots, x_n))) = \text{etc.}) \end{aligned}$$

If X_1, X_2, \dots, X_n are classes, their *Cartesian product* is the class:

$$X_1 \times X_2 \times \dots \times X_n = \{(x_1, \dots, x_n) \mid x_1 \in X_1 \wedge \dots \wedge x_n \in X_n\}.$$

We write X^2 in place of $X \times X$, X^3 for $X \times X \times X$, etc.

Let X be any class. An n -ary relation on X is a class $R \subseteq X^n$. We often write $R(\vec{x})$ in place of $(\vec{x}) \in R$.

Suppose R is an $(n + 1)$ -ary relation on a class X , where $n > 0$. The *domain* of R is the class

$$\text{dom}(R) = \{(\vec{x}) \mid \exists y R(y, \vec{x})\}.$$

The *range* of R is the class

$$\text{ran}(R) = \{y \mid \exists \vec{x} R(y, \vec{x})\}.$$

If $Z \subseteq X$, we set

$$R \upharpoonright Z = \{(y, \vec{x}) \in R \mid \vec{x} \in Z\}.$$

(Notice that according to our conventions concerning finite strings of variables, $\vec{x} \in Z$ means $x_1 \in Z, \dots, x_n \in Z$. If we want to mean that $(x_1, \dots, x_n) \in Z$ we would write $(\vec{x}) \in Z$.)

We define

$$R''Z = \text{ran}(R \upharpoonright Z).$$

Let X be a class, $n > 0$. An n -ary function over X is an $(n + 1)$ -ary relation R on X such that

$$(\forall (\vec{x}) \in \text{dom}(R))(\exists! y) R(y, \vec{x}).$$

We often write $R(\vec{x}) = y$ instead of $R(y, \vec{x})$ in such cases. Thus $R(\vec{x})$ is the unique y such that $R(y, \vec{x})$. We say that R is *total* on X iff $\text{dom}(R) = X$.

Let f be an n -ary function over V . We write

$$f: X \rightarrow Y$$

to denote that (f is a function and) $\text{dom}(f) = X$ and $\text{ran}(f) \subseteq Y$. We say that f is *one-one* (or *injective*), and write

$$f: X \xrightarrow{(1-1)} Y,$$

iff for all $x_1, x_2 \in X$,

$$x_1 \neq x_2 \rightarrow f(x_1) \neq f(x_2).$$

We say f is *onto* Y (or is a *surjection* to Y), and write

$$f: X \xrightarrow{\text{onto}} Y,$$

iff $\text{ran}(f) = Y$. We say f is a *bijection* iff it is both one-one and onto, and write

$$f: X \leftrightarrow Y.$$

If f is bijective there is a unique function $f^{-1}: Y \rightarrow X$ (called the *inverse* of f) such that

$$(\forall x \in X)(f^{-1}(f(x)) = x),$$

$$(\forall y \in Y)(f(f^{-1}(y)) = y).$$

Regardless of whether or not f is bijective, we set, for any $Z \subseteq Y$,

$$f^{-1}Z = \{x \in X \mid f(x) \in Z\}.$$

The set $f^{-1}Z$ is called the *preimage of Z under f* .

Notice that by our definition of the n -tuples (x_1, \dots, x_n) , every function is a set of ordered pairs, regardless of whether or not the function is unary.

If X and Y are structures of the same type, we write

$$f: X \cong Y$$

if f is a bijection from X to Y which preserves the structure (i.e. if f is an *isomorphism*).

We denote the composition of functions f, g by $f \circ g$, as usual. Thus if $f: Y \rightarrow Z$ and $g: X \rightarrow Y$, we define $f \circ g: X \rightarrow Z$ by

$$(\forall x \in X)(f \circ g(x) = f(g(x))).$$

For any sets x, y we define

$${}^x y = \{f \mid f: x \rightarrow y\}.$$

The *identity function* is the unary function

$$\text{id} = \{(y, x) \mid y = x\}.$$

Of course, being a proper class, id is not strictly speaking a function at all, but for any set X , $\text{id} \upharpoonright X$ will be a function, so this definition is convenient.

A function whose domain is an ordinal is called a *sequence*; if α is that ordinal domain, we say that the sequence is an α -sequence. If f is an α -sequence, and if $f(\xi) = x_\xi$ for all $\xi < \alpha$, we often write $f = (x_\xi \mid \xi < \alpha)$.

If $f: I \rightarrow V$, and if we denote $f(i)$ by x_i for each $i \in I$, we often write $\{x_i \mid i \in I\}$ in place of $f'' I$, $\bigcup_{i \in I} x_i$ in place of $\cup (f'' I)$, and $\bigcap_{i \in I} x_i$ in place of $\cap (f'' I)$. Similarly, given a sequence $f = (x_v \mid v < \tau)$, we sometimes write $\bigcup_{v < \tau} x_v$ for $\cup (f'' \tau)$ and $\bigcap_{v < \tau} x_v$ for $\cap (f'' \tau)$. And if $h = (\alpha_v \mid v < \tau)$ is a sequence of ordinals, we would write $\sup_{v < \tau} \alpha_v$ for $\sup (h'' \tau)$.

The inverse functions to the ordered pair function are defined thus:

if $u = (x, y)$, then $(u)_0 = x$ and $(u)_1 = y$;

if u is not an ordered pair, then $(u)_0 = (u)_1 = \emptyset$.

Similarly we define inverse functions $(u)_0^n, \dots, (u)_{n-1}^n$ to the n -tuple function.

3.4 (Induction and Recursion). By using the Axiom of Foundation (together with other axioms of ZF), every instance of the following *schema of proof by \in -induction* can be proved in ZF:

$$\forall x [(\forall y \in x) \Phi(y) \rightarrow \Phi(x)] \rightarrow \forall x \Phi(x).$$

More generally, if X is any class, a relation $R \subseteq X^2$ is said to be *well-founded* iff:

- (i) $(\forall x \in \text{dom}(R)) [\{y \mid R(y, x)\}$ is a set];
- (ii) $\forall a [a \neq \emptyset \wedge a \subseteq X \rightarrow (\exists x \in a) (\forall y \in a) \neg R(y, x)]$.

If R is such a relation, then every instance of the following *schema of proof by induction on R* is a theorem of ZF:

$$(\forall x \in X) [(\forall y \in X) (R(y, x) \rightarrow \Phi(y)) \rightarrow \Phi(x)] \rightarrow (\forall x \in X) \Phi(x).$$

It follows from the above that every instance of the following *schema of definition by recursion* is provable in ZF. Let G be a total $(n + 2)$ -ary function over V , and let H be a total unary function over V such that the relation $\{(z, y) \mid z \in H(y)\}$ is well-founded. Then there is a unique, total $(n + 1)$ -ary function F over V such that

$$F(y, \vec{a}) = G(y, \vec{a}, F \upharpoonright (H(y) \times \{(\vec{a})\}))$$

(Actually, some care is required in formulating this result precisely. Given formulas which determine G and H , possibly with reference to certain set parameters, one can explicitly write down a third formula such that in ZF it is provable that the class determined by this formula has all of the properties required of F above. We assume the reader is quite familiar with all of this, though in fact we shall not really need to know the exact formulation. As is usually the case in set theory, all that we require is the knowledge that ZF allows definitions “of a recursive nature”).

A particular case of the above recursion principle is when $H = \text{id}$, when it is called the *principle of \in -recursion*. There are also recursion principles applicable to functions defined not on all of V but on some class, the most common example being when the class concerned is the class of all ordinals; \in -recursion restricted to the ordinals is known as *ordinal recursion*, and will be used frequently in this book.

The total, unary function TC (*transitive closure*) is defined by the \in -recursion

$$\text{TC}(x) = x \cup \bigcup \{\text{TC}(y) \mid y \in x\}.$$

(Intuitively, $\text{TC}(x) = x \cup (\bigcup x) \cup (\bigcup \bigcup x) \cup (\bigcup \bigcup \bigcup x) \cup \dots$. It is not hard to show that $\text{TC}(x)$ is the \subseteq -smallest transitive set y such that $x \subseteq y$. We call the set $\text{TC}(x)$ the *transitive closure* of x .

The relation $\{(x, y) \mid x \in \text{TC}(y)\}$ is well-founded. Hence we can carry out definitions by recursion on this relation. This form of recursive definition will also be quite common in this book.

3.5 (The Cumulative Hierarchy). The *cumulative hierarchy of sets* is defined by the ordinal recursion

$$\begin{aligned} V_0 &= \emptyset; \\ V_\alpha &= \bigcup \{\mathcal{P}(V_\beta) \mid \beta < \alpha\}. \end{aligned}$$

It is not hard to see that

$$V = \bigcup \{V_\alpha \mid \alpha \in \text{On}\}.$$

(The proof makes central use of the Axiom of Foundation, and indeed the above “equation” may be taken as an alternative formulation of this axiom.)

The *rank* of a set x is the least ordinal α such that $x \in V_{\alpha+1}$. We may define the rank function directly by means of the \in -recursion

$$\begin{aligned} \text{rank}(\emptyset) &= 0 \\ \text{rank}(x) &= \bigcup \{\text{rank}(y) + 1 \mid y \in x\}. \end{aligned}$$

Notice that $x \in y$ implies $\text{rank}(x) < \text{rank}(y)$.

4. Ordinal Numbers

The notion of an ordinal number plays a central role in set theory, and we have referred to ordinals several times already. In this section we consider, very briefly, the arithmetic of ordinal numbers.

Let α, β be ordinals. We define the *ordinal sum* $\alpha + \beta$ by recursion on β , thus:

$$\begin{aligned}\alpha + 0 &= \alpha; \\ \alpha + (\beta + 1) &= (\alpha + \beta) + 1; \\ \alpha + \delta &= \bigcup \{(\alpha + \beta) \mid \beta < \delta\}, \quad \text{if } \lim(\delta).\end{aligned}$$

The ordinal sum is not commutative; for example, $1 + \omega = \omega$ but $\omega + 1 > \omega$.

The *ordinal product* $\alpha \cdot \beta$ is defined by the following recursion on β :

$$\begin{aligned}\alpha \cdot 0 &= 0; \\ \alpha \cdot (\beta + 1) &= (\alpha \cdot \beta) + \alpha; \\ \alpha \cdot \delta &= \bigcup \{\alpha \cdot \beta \mid \beta < \delta\}, \quad \text{if } \lim(\delta).\end{aligned}$$

The ordinal product is not commutative; for example $2 \cdot \omega = \omega$ but $\omega \cdot 2 = \omega + \omega > \omega$.

(Both ordinal sum and ordinal product can be defined in an alternative fashion, but we shall not go into that here.)

Notice that $(\omega \cdot \alpha \mid \alpha \in \text{On})$ enumerates 0 and all the limit ordinals.

For $\alpha > 0$, the *ordinal power* α^β is defined by the following recursion on β :

$$\begin{aligned}\alpha^0 &= 1; \\ \alpha^{(\beta+1)} &= (\alpha^\beta) \cdot \alpha; \\ \alpha^\delta &= \bigcup \{\alpha^\beta \mid \beta < \delta\}, \quad \text{if } \lim(\delta).\end{aligned}$$

We shall not be concerned with any of the properties of ordinal exponentiation.

5. Cardinal Numbers

A *cardinal number* (or simply, a *cardinal*) is an ordinal α such that there is no $\beta < \alpha$ for which there is a function $f: \beta \xrightarrow{\text{onto}} \alpha$.

Clearly, $0, 1, 2, 3, \dots, n, \dots, \omega$ are all cardinals. All other cardinals are said to be *uncountable*. We generally use κ, λ, μ , to denote cardinals.

Using AC it can be shown that for every set x there is a unique cardinal κ for which there is a bijection $f: \kappa \leftrightarrow x$. We call κ the *cardinality* of x , denoted by $|x|$.

Clearly, if κ is a cardinal, then, recalling that κ is the set $\{\alpha \mid \alpha < \kappa\}$, we have $|\kappa| = \kappa$.

If κ is a cardinal, the least cardinal greater than κ is called the (*cardinal*) *successor* of κ , and is denoted by κ^+ . For convenience, we extend this notation so that for any *ordinal* α , α^+ denotes the least *cardinal* greater than α . (So in particular, we have $\alpha^+ = |\alpha|^+$.) Any infinite cardinal of the form κ^+ is called a *successor*

cardinal. An infinite cardinal which is not a successor cardinal is called a *limit cardinal*.

Clearly, if $\lim(\beta)$ and $(\kappa_\alpha \mid \alpha < \beta)$ is a strictly increasing sequence of cardinals, then $\sup_{\alpha < \beta} \kappa_\alpha$ is a limit cardinal.

The canonical, monotone enumeration of the infinite cardinals, $(\omega_\alpha \mid \alpha \in \text{On})$, is defined by the following recursion:

$$\begin{aligned}\omega_0 &= \omega; \\ \omega_{\alpha+1} &= \omega_\alpha^+; \\ \omega_\delta &= \sup_{\alpha < \delta} \omega_\alpha, \quad \text{if } \lim(\delta).\end{aligned}$$

Of course, each cardinal ω_α is also an ordinal. In order to distinguish between the two cases when ω_α is being used as a cardinal and when it is being used as an ordinal, many texts use the symbol \aleph_α (“aleph- α ”) to denote ω_α regarded as a cardinal, and reserve the notation ω_α for pure ordinal use. However, in this book we shall have very little occasion to use ω_α as an ordinal (in the strict sense), so we shall rely upon the single notation ω_α in all cases.

Notice that ω_α is a limit cardinal iff $\alpha = 0$ or $\lim(\alpha)$.

We write

$$\forall \kappa \Phi(\kappa)$$

in place of

$$\forall \alpha [\alpha \text{ is a cardinal} \rightarrow \Phi(\alpha)].$$

and

$$\exists \kappa \Phi(\kappa)$$

in place of

$$\exists \alpha [\alpha \text{ is a cardinal} \wedge \Phi(\alpha)].$$

Let $(\kappa_\alpha \mid \alpha < \beta)$ be a sequence of cardinals. The *cardinal sum* $\sum_{\alpha < \beta} \kappa_\alpha$ is defined thus:

$$\sum_{\alpha < \beta} \kappa_\alpha = |\{(\xi, \alpha) \mid \xi < \kappa_\alpha \wedge \alpha < \beta\}|.$$

Clearly, $\sum_{\alpha < \beta} \kappa_\alpha$ is the cardinality of the union of any disjoint collection $\{A_\alpha \mid \alpha < \beta\}$ of sets such that $|A_\alpha| = \kappa_\alpha$.

Clearly,

$$\sum_{\alpha < \beta} \kappa_\alpha = \sup_{\alpha < \beta} \kappa_\alpha.$$

We write $\kappa_0 + \kappa_1$ instead of $\sum_{\alpha < 2} \kappa_\alpha$. Provided that at least one of κ_0, κ_1 is infinite, we have

$$\kappa_0 + \kappa_1 = \max(\kappa_0, \kappa_1).$$

We usually rely upon context to distinguish between ordinal and cardinal addition, rather than introduce additional notation. For instance, $\alpha + \beta$ would usually mean ordinal addition, whereas $\kappa + \lambda$ would mean cardinal addition.

The *cardinal product*, $\prod_{\alpha < \beta} \kappa_\alpha$, is defined thus:

$$\prod_{\alpha < \beta} \kappa_\alpha = |\{f \mid f: \beta \rightarrow \bigcup_{\alpha < \beta} \kappa_\alpha \wedge (\forall \alpha < \beta)(f(\alpha) \in \kappa_\alpha)\}|.$$

Clearly, $\prod_{\alpha < \beta} \kappa_\alpha$ is the cardinality of the *cartesian product*

$$\prod_{\alpha < \beta} A_\alpha = \{f \mid f: \beta \rightarrow \bigcup_{\alpha < \beta} A_\alpha \ \& \ (\forall \alpha < \beta)(f(\alpha) \in A_\alpha)\}$$

of any family of sets A_α , $\alpha < \beta$, such that $|A_\alpha| = \kappa_\alpha$.

We write $\kappa_0 \cdot \kappa_1$ instead of $\prod_{\alpha < 2} \kappa_\alpha$. If at least one of κ_0, κ_1 is infinite and neither is 0, then

$$\kappa_0 \cdot \kappa_1 = \max(\kappa_0, \kappa_1).$$

The *cardinal power*, κ^λ , is defined by

$$\kappa^\lambda = \prod_{\alpha < \lambda} \kappa.$$

(Again, context and notation are used to distinguish between cardinal and ordinal exponentiation.) Recalling that for any sets x, y ,

$${}^x y = \{f \mid f: x \rightarrow y\},$$

we see that if $|x| = \lambda$ and $|y| = \kappa$ then

$$\kappa^\lambda = |{}^x y|.$$

In particular,

$$\kappa^\lambda = |{}^\lambda \kappa| = |\{f \mid f: \lambda \rightarrow \kappa\}|.$$

By considering characteristic functions of subsets of x , we see easily that

$$|\mathcal{P}(x)| = 2^{|x|},$$

for any set x . Consequently, for any cardinal κ ,

$$2^\kappa = |\mathcal{P}(\kappa)|.$$

By means of the well-known Cantor diagonal argument, it follows that

$$2^\kappa > \kappa.$$

Hence

$$2^\kappa \geq \kappa^+.$$

For finite κ , we only have equality in the above in the cases $\kappa = 0$ and $\kappa = 1$. But for infinite κ , the axioms of ZFC set theory do not provide enough information to decide whether or not $2^\kappa = \kappa^+$. (The precise situation is rather complicated, and certainly outside the scope of this book.) The statement

$$2^\omega = \omega_1,$$

which is neither provable nor refutable in ZFC, is known as the *continuum hypothesis* (CH). The statement

$$\forall \alpha (2^{\omega_\alpha} = \omega_{\alpha+1}),$$

which is likewise neither provable nor refutable in ZFC, is known as the *generalised continuum hypothesis* (GCH). (The word “continuum” is used here because 2^ω is the cardinality of the real number continuum.)

5.1 Lemma.

- (i) \sum and \prod are commutative and associative operations on cardinal numbers.
(ii) \prod distributes over \sum ; i.e.

$$\prod_{\alpha < \gamma} \sum_{\beta < \delta} \kappa_{\alpha\beta} = \sum_{f \in {}^\gamma \delta} \prod_{\alpha < \gamma} \kappa_{\alpha f(\alpha)}.$$

$$(iii) \quad \kappa \cdot \sum_{\alpha < \beta} \kappa_\alpha = \sum_{\alpha < \beta} \kappa \cdot \kappa_\alpha.$$

$$(iv) \quad \sum_{\alpha < \beta} \kappa = |\beta| \cdot \kappa.$$

Proof. An easy exercise. \square

5.2 Lemma. If $2 \leq \kappa \leq \lambda$ and $\lambda \geq \omega$, then $\kappa^\lambda = 2^\lambda$.

Proof. Clearly, $2^\lambda \leq \kappa^\lambda$. Conversely,

$$\kappa^\lambda \leq \lambda^\lambda \leq (2^\lambda)^\lambda = 2^{\lambda \cdot \lambda} = 2^\lambda. \quad \square$$

5.3 Lemma (“The König Inequality”). If $\kappa_\alpha < \lambda_\alpha$ for all $\alpha < \beta$, then

$$\sum_{\alpha < \beta} \kappa_\alpha < \prod_{\alpha < \beta} \lambda_\alpha.$$

Proof. Define

$$f: \bigcup_{\alpha < \beta} (\kappa_\alpha \times \{\alpha\}) \rightarrow \prod_{\alpha < \beta} \lambda_\alpha$$

by setting

$$[f(\xi, \alpha)](v) = \begin{cases} \xi + 1, & \text{if } v = \alpha \\ 0, & \text{if } v \neq \alpha \end{cases}$$

Clearly, f is one-one. Hence

$$\sum_{\alpha < \beta} \kappa_\alpha \leq \prod_{\alpha < \beta} \lambda_\alpha.$$

We show that equality is impossible. Let

$$h: \bigcup_{\alpha < \beta} (\kappa_\alpha \times \{\alpha\}) \rightarrow \prod_{\alpha < \beta} \lambda_\alpha.$$

We show that h cannot be onto. For $\gamma < \beta$, define

$$h_\gamma: \bigcup_{\alpha < \beta} (\kappa_\alpha \times \{\alpha\}) \rightarrow \lambda_\gamma$$

by

$$h_\gamma(\xi, \alpha) = [h(\xi, \alpha)](\gamma).$$

Since $\kappa_\gamma < \lambda_\gamma$, $h_\gamma \upharpoonright (\kappa_\gamma \times \{\gamma\})$ cannot map onto λ_γ , so we can pick

$$a_\gamma \in \lambda_\gamma - h_\gamma''(\kappa_\gamma \times \{\gamma\}).$$

Define $g \in \prod_{\alpha < \beta} \lambda_\alpha$ by

$$g(\gamma) = a_\gamma \quad (\gamma < \beta).$$

Clearly, $g \notin \text{ran}(h)$, so h is not onto, and we are done. \square

5.4 Lemma. Let κ_α be cardinals for $\alpha < \beta$, and set

$$\kappa = \sum_{\alpha < \beta} \kappa_\alpha.$$

For any cardinal λ ,

$$\lambda^\kappa = \prod_{\alpha < \beta} \lambda^{\kappa_\alpha}.$$

Proof. Let

$$X = \bigcup_{\alpha < \beta} (\kappa_\alpha \times \{\alpha\}).$$

Thus $|X| = \kappa$. Let $f \in {}^X \lambda$. For $\alpha < \beta$, define $f_\alpha \in {}^{(\kappa_\alpha)} \lambda$ by

$$f_\alpha(v) = f(v, \alpha).$$

Then

$$(f_\alpha \mid \alpha < \beta) \in \prod_{\alpha < \beta} {}^{(\kappa_\alpha)} \lambda.$$

Since the mapping

$$f \mapsto (f_\alpha \mid \alpha < \beta)$$

from ${}^X \lambda$ to $\prod_{\alpha < \beta} {}^{(\kappa_\alpha)} \lambda$ is clearly one-one and onto, the desired equality is proved. \square

A subset A of a limit² ordinal α is said to be *unbounded* in α iff for no $\gamma < \alpha$ do we have $A \subseteq \gamma$. Equivalently, $A \subseteq \alpha$ is unbounded in α iff

$$(\forall v \in \alpha)(\exists \tau \in A)(\tau \geq v).$$

(i.e. iff $\sup(A) = \alpha$.)

Let γ, α be limit ordinals. A function $f: \gamma \rightarrow \alpha$ is said to be *cofinal* iff f is order-preserving and $\text{ran}(f)$ is an unbounded subset of α .

Let α be a limit ordinal. The *cofinality* of α is the least ordinal γ such that there is a cofinal function $f: \gamma \rightarrow \alpha$. We denote the cofinality of α by $\text{cf}(\alpha)$. It is easily seen that $\text{cf}(\alpha)$ is always a cardinal.

A limit ordinal α is said to be *regular* iff $\text{cf}(\alpha) = \alpha$; otherwise it is *singular*. Every regular ordinal is a cardinal. Also, $\text{cf}(\alpha)$ is always a regular cardinal. The cardinal ω is regular; the cardinal ω_ω is singular of cofinality ω .

5.5 Lemma. *Let κ be an infinite cardinal. Then $\text{cf}(\kappa)$ is the least α such that there are cardinals $\kappa_\xi < \kappa$, $\xi < \alpha$ such that*

$$\kappa = \sum_{\xi < \alpha} \kappa_\xi.$$

Proof. Let $\lambda = \text{cf}(\kappa)$, and let α be least such that $\kappa = \sum_{\xi < \alpha} \kappa_\xi$ for some $\kappa_\xi < \kappa$. We must show that $\lambda = \alpha$.

Let $(\gamma_\xi \mid \xi < \lambda)$ be cofinal in κ . For each $\xi < \lambda$, $|\gamma_\xi| < \kappa$. But $\kappa = \bigcup_{\xi < \lambda} \gamma_\xi$. It follows easily that $\kappa = \sum_{\xi < \lambda} |\gamma_\xi|$. Hence $\alpha \leq \lambda$.

Suppose that $\alpha < \lambda$. Pick $\kappa_\xi < \kappa$ for $\xi < \alpha$ so that $\kappa = \sum_{\xi < \alpha} \kappa_\xi$. Since $\alpha < \lambda$, $(\kappa_\xi \mid \xi < \alpha)$ is not cofinal in κ , so for some $\gamma < \kappa$, we have $\kappa_\xi \leq \gamma$ for all $\xi < \alpha$. Hence $\sum_{\xi < \alpha} \kappa_\xi \leq \sum_{\xi < \alpha} |\gamma| = |\alpha| \cdot |\gamma| < \kappa$, which is a contradiction. Thus $\alpha = \lambda$. \square

² On a formal level, this and the following definitions can be applied to any ordinal, but the notions are trivial in the case of successor ordinals.

5.6 Lemma. *Let κ be an infinite cardinal. Then κ is regular iff*

$$(\forall \lambda < \kappa)({}^\lambda \kappa = \bigcup_{\alpha < \kappa} {}^\lambda \alpha).$$

Proof. (\rightarrow) Let $\lambda < \kappa$. If $f \in \bigcup_{\alpha < \kappa} {}^\lambda \alpha$, then $f \in {}^\lambda \kappa$. Conversely, if $f \in {}^\lambda \kappa$, then since $\text{cf}(\kappa) = \kappa > \lambda$, $\text{ran}(f) \subseteq \alpha$ for some $\alpha < \kappa$, so $f \in {}^\lambda \alpha$. Thus ${}^\lambda \kappa = \bigcup_{\alpha < \kappa} {}^\lambda \alpha$.

(\leftarrow) Let $\lambda < \kappa$, $f: \lambda \rightarrow \kappa$. For some $\alpha < \kappa$, $f \in {}^\lambda \alpha$, so f cannot be cofinal. Hence κ is regular. \square

5.7 Lemma. *Let κ be an infinite cardinal. Then κ^+ is regular.*

Proof. Suppose $\text{cf}(\kappa^+) \leq \kappa$. By 5.5 there are cardinals $\kappa_\alpha < \kappa^+$ such that

$$\kappa^+ \leq \sum_{\alpha < \kappa} \kappa_\alpha.$$

Then

$$\kappa^+ \leq \sum_{\alpha < \kappa} \kappa_\alpha \leq \sum_{\alpha < \kappa} \kappa = \kappa \cdot \kappa = \kappa,$$

a contradiction. \square

5.8 Lemma. *Let κ be an infinite cardinal. Then $\kappa^{\text{cf}(\kappa)} > \kappa$.*

Proof. Let $\kappa = \sum_{\alpha < \text{cf}(\kappa)} \kappa_\alpha$, where $\kappa_\alpha < \kappa$. By 5.3 we have

$$\kappa = \sum_{\alpha < \text{cf}(\kappa)} \kappa_\alpha < \prod_{\alpha < \text{cf}(\kappa)} \kappa = \kappa^{\text{cf}(\kappa)}. \quad \square$$

The following result shows that under certain circumstances $\text{cf}(\kappa)$ may be the least cardinal λ such that $\kappa^\lambda > \kappa$.

5.9 Lemma. *Let κ be an infinite cardinal. If $\lambda < \text{cf}(\kappa)$ and $(\forall \mu < \kappa)(2^\mu \leq \kappa)$, then $\kappa^\lambda = \kappa$.*

Proof. By an argument as in 5.6 we see that if $\lambda < \text{cf}(\kappa)$,

$$\kappa^\lambda = |{}^\lambda \kappa| = \left| \bigcup_{\alpha < \kappa} {}^\lambda \alpha \right| \leq \sum_{\alpha < \kappa} |\alpha|^\lambda.$$

But for $\alpha < \kappa$ we have

$$|\alpha|^\lambda \leq (2^{|\alpha|})^\lambda = 2^{|\alpha| \cdot \lambda} \leq \kappa.$$

Hence

$$\kappa \leq \kappa^\lambda \leq \kappa. \quad \square$$

We define the *weak power* of κ by λ thus:

$$\kappa^{<\lambda} = \sum_{\mu < \lambda} \kappa^\mu.$$

5.10 Lemma. *Let κ be an infinite cardinal.*

- (i) *If κ is regular and $(\forall \mu < \kappa)(2^\mu \leq \kappa)$, then $\kappa^{<\kappa} = \kappa$.*
- (ii) *Assume GCH. Then κ is regular iff $\kappa^{<\kappa} = \kappa$.*

Proof. (i) By 5.9, $\kappa^\lambda = \kappa$ for all $\lambda < \text{cf}(\kappa) = \kappa$. Thus

$$\kappa^{<\kappa} = \sum_{\lambda < \kappa} \kappa^\lambda = \sum_{\lambda < \kappa} \kappa = \kappa.$$

(ii) If κ is regular then (since GCH implies $2^\mu = \mu^+ \leq \kappa$ for all $\mu < \kappa$) by (i) we have $\kappa^{<\kappa} = \kappa$. If κ is singular, then by 5.8, $\kappa^{<\kappa} \geq \kappa^{\text{cf}(\kappa)} > \kappa$. \square

5.11 Lemma. *Let κ be an infinite cardinal.*

- (i) $\kappa^\kappa \geq \kappa^{<\kappa} \geq 2^{<\kappa} \geq \kappa$.
- (ii) *If $\kappa = \lambda^+$, then $\kappa^{<\kappa} = 2^{<\kappa} = 2^\lambda$.*

Proof. (i) The only non-trivial inequality is $2^{<\kappa} \geq \kappa$. And this follows quite easily from the fact that for every $\lambda < \kappa$, $2^{<\kappa} \geq 2^\lambda > \lambda$.

(ii) We have $2^{<\kappa} = 2^\lambda$ and $\kappa^{<\kappa} = \kappa^\lambda$. Hence

$$2^\lambda = 2^{<\kappa} \leq \kappa^{<\kappa} = \kappa^\lambda \leq (2^\lambda)^\lambda = 2^\lambda. \quad \square$$

5.12 Lemma. $\text{GCH} \leftrightarrow (\forall \kappa \geq \omega)(2^{<\kappa} = \kappa)$.

Proof. (\rightarrow) By GCH,

$$2^{<\kappa} = \sum_{\lambda < \kappa} 2^\lambda = \sum_{\lambda < \kappa} \lambda^+ = \kappa,$$

for any infinite κ .

(\leftarrow) For any infinite κ ,

$$2^\kappa = 2^{<\kappa^+} = \kappa^+. \quad \square$$

6. Closed Unbounded Sets

Let α be a limit ordinal. Recall that a set $A \subseteq \alpha$ is *unbounded* in α iff

$$(\forall v \in \alpha)(\exists \tau \in A)(\tau \geq v).$$

A set $A \subseteq \alpha$ is *closed* in α iff $\bigcup(A \cap \gamma) \in A$ for all $\gamma < \alpha$. Equivalently, if we define a *limit point* of A to be any limit ordinal γ such that $A \cap \gamma$ is unbounded in γ , then A will be closed in α iff it contains all its limit points below α . Still another formulation is that A is closed in α iff, whenever $\lim(\tau)$ and $(\alpha_v \mid v < \tau)$ is a strictly increasing sequence of elements of A which is not cofinal in α , then $\bigcup_{v < \tau} \alpha_v \in A$.

We use the abbreviation “club” to mean “closed and unbounded”. Club sets play an important role in our development.

6.1 Lemma. *Let κ be an infinite cardinal, $\text{cf}(\kappa) > \omega$. If A, B are club subsets of κ , then $A \cap B$ is club in κ .*

Proof. That $A \cap B$ is closed in κ is obvious. To establish unboundedness, let $\alpha < \kappa$ be given. Pick $\beta_0 \in A, \beta_0 > \alpha$. By recursion now, let β_{2n+1} be the least member of B than β_{2n} and let β_{2n+2} be the least member of A greater than β_{2n+1} . Let $\beta = \bigcup_{n < \omega} \beta_n$. Since $\beta = \bigcup_{n < \omega} \beta_{2n}$ and A is closed we have $\beta \in A$, and similiary $\beta = \bigcup_{n < \omega} \beta_{2n+1}$ implies $\beta \in B$. Thus $\beta \in A \cap B, \beta > \alpha$, and we are done. \square

By generalising the above proof we obtain:

6.2 Lemma. *Let κ be an uncountable regular cardinal. If $\lambda < \kappa$ and $A_\nu, \nu < \lambda$, are club subsets of κ , then $\bigcap_{\nu < \lambda} A_\nu$ is a club subset of κ . \square*

Let α be an infinite ordinal. A non-decreasing function $f: \alpha \rightarrow \text{On}$ is said to be *continuous* if for every limit ordinal $\delta < \alpha$,

$$f(\delta) = \bigcup_{\beta < \delta} f(\beta).$$

A *normal function* on α is a (strictly) increasing, continuous function $f: \alpha \rightarrow \text{On}$.

6.3 Lemma. *Let α be a limit ordinal. If f is an increasing function from α into On , then $f(\gamma) \geq \gamma$ for all $\gamma \in \alpha$.*

Proof. By induction on γ . If the result holds below γ , then for all $\beta < \gamma$ we have $f(\gamma) > f(\beta) \geq \beta$, and hence $f(\gamma) \geq \gamma$. \square

6.4 Lemma. *Let κ be an uncountable regular cardinal.*

(i) *If $A \subseteq \kappa$ is club, then the enumeration of A in increasing order (as ordinals) is a normal function from κ to κ .*

(ii) *If $f: \kappa \rightarrow \kappa$ is a normal function, then $\text{ran}(f)$ is a club subset of κ .*

Proof. Trivial. \square

6.5 Lemma. *Let κ be an uncountable regular cardinal, f a normal function from κ to κ . Then the set $\{\alpha \in \kappa \mid f(\alpha) = \alpha\}$ is club in κ .*

Proof. Let

$$A = \{\alpha \in \kappa \mid f(\alpha) = \alpha\}.$$

If $\gamma < \kappa$ is a limit point of A , then

$$f(\gamma) = \bigcup_{\alpha < \gamma} f(\alpha) = \bigcup_{\alpha \in A \cap \gamma} f(\alpha) = \bigcup_{\alpha \in A \cap \gamma} \alpha = \gamma$$

(since $\bigcup (A \cap \gamma) = \gamma$), so $\gamma \in A$. Hence A is closed in κ .

To show that A is unbounded in κ , let $\alpha_0 \in \kappa$ be given. By recursion, define $\alpha_{n+1} = f(\alpha_n)$. Set $\alpha = \bigcup_{n < \omega} \alpha_n$. Then by 6.3, we have $\alpha \geq \alpha_0$ and, by continuity,

$$f(\alpha) = \bigcup_{n < \omega} f(\alpha_n) = \bigcup_{n < \omega} \alpha_{n+1} = \alpha,$$

so $\alpha \in A$. Hence A is unbounded in κ . \square

6.6 Lemma. *Let κ be an uncountable regular cardinal, and let $h: \kappa \rightarrow \kappa$. Set*

$$A = \{\gamma \in \kappa \mid (\forall v < \gamma)(h(v) < \gamma)\}.$$

Then A is club in κ .

Proof. It is immediate that A is closed in κ . To prove unboundedness, given $\alpha_0 \in \kappa$, define, by recursion, α_{n+1} to be the least γ such that $h''\alpha_n \subseteq \gamma$, and set $\alpha = \bigcup_{n < \omega} \alpha_n$. Clearly, $\alpha \geq \alpha_0$ and $\alpha \in A$. \square

7. The Collapsing Lemma

In this section we prove a simple lemma which is of extreme importance in constructibility theory: the Mostowski-Shepherdson Collapsing Lemma. We start with a definition.

A set X is said to be *extensional* if:

$$(\forall u, v \in X)(u \neq v \rightarrow (\exists x \in X)(x \in u \leftrightarrow x \notin v))$$

(In other words, a set X is extensional iff the structure $\langle X, \in \rangle$ is a model of the Axiom of Extensionality.)

7.1 Theorem (The Collapsing Lemma). *Let X be an extensional set. Then there is a unique transitive set M and a unique bijection $\pi: X \leftrightarrow M$ such that*

$$\pi: \langle X, \in \rangle \cong \langle M, \in \rangle.$$

*Moreover, if $Y \subseteq X$ is transitive, then $\pi \upharpoonright Y = \text{id} \upharpoonright Y$. (The transitive set M is called the *transitive collapse* or *transitivisation* of X .)*

Proof. We first of all see what a function π as in the theorem must look like. This will amount to a proof of the uniqueness of π and M . So suppose that

$$\pi: \langle X, \in \rangle \cong \langle M, \in \rangle,$$

where M is transitive. Let $x, y \in X$, $y \in x$. By isomorphism, $\pi(y) \in \pi(x)$. Hence

$$\{\pi(y) \mid y \in X \wedge y \in x\} \subseteq \pi(x).$$

Now let $x \in X$, $z \in \pi(x)$. Since M is transitive, $z \in M$. So for some $y \in X$, $z = \pi(y)$. Then $\pi(y) \in \pi(x)$, so as π is an isomorphism, $y \in x$. Thus

$$\pi(x) \subseteq \{\pi(y) \mid y \in X \wedge y \in x\}.$$

Hence

$$\pi(x) = \{\pi(y) \mid y \in X \wedge y \in x\}.$$

This shows us what form π must have, as well as providing a proof of the uniqueness of such a π (and hence also of $M = \text{ran}(\pi)$). (More precisely, since π , M necessarily have the above structure, given any two candidates π_1, M_1 and π_2, M_2 , a trivial \in -induction shows that $\pi_1(x) = \pi_2(x)$ for all $x \in X$, so $\pi_1 = \pi_2$ and $M_1 = M_2$.)

We prove existence. Define π on X by the \in -recursion:

$$\pi(x) = \{\pi(y) \mid y \in X, y \in x\}.$$

Set $M = \text{ran}(\pi)$. We show that π, M are as required.

We prove first that π is one-one. Suppose not, and pick $x_1 \in X$ of least rank such that for some $x_2 \in X$, $x_2 \neq x_1$ and $\pi(x_2) = \pi(x_1)$. Since X is extensional there is a $y \in X$ such that either $y \in x_1$, $y \notin x_2$ or else $y \notin x_1$, $y \in x_2$.

Suppose first that $y \in x_1$, $y \notin x_2$. By definition of $\pi(x_1)$ we have $\pi(y) \in \pi(x_1)$. Hence as $\pi(x_1) = \pi(x_2)$, we have $\pi(y) \in \pi(x_2)$. So by definition of $\pi(x_2)$ there must be a $z \in X$, $z \in x_2$, such that $\pi(y) = \pi(z)$. Since $y \notin x_2$, $y \neq z$. But $\text{rank}(y) < \text{rank}(x_1)$ so the existence of such a z contradicts the choice of x_1 .

Now suppose that $y \notin x_1$, $y \in x_2$. By definition of $\pi(x_2)$, we have $\pi(y) \in \pi(x_2)$. Hence as $\pi(x_1) = \pi(x_2)$, we have $\pi(y) \in \pi(x_1)$. So by definition of $\pi(x_1)$ there must be a $z \in X$, $z \in x_1$, such that $\pi(y) = \pi(z)$. Since $y \notin x_1$, $y \neq z$. But $\text{rank}(z) < \text{rank}(x_1)$, so the existence of such a y (for z) contradicts the choice of x_1 .

Having established that π is one-one, we show next that π is an \in -isomorphism. Let $x, y \in X$. If $x \in y$, then by definition of $\pi(y)$ we have $\pi(x) \in \pi(y)$. Conversely, suppose $\pi(x) \in \pi(y)$. Then by definition of $\pi(y)$, $\pi(x) = \pi(z)$ for some $z \in X$, $z \in y$. But π is one-one. Hence $x = z$, giving $x \in y$. Thus π is an \in -isomorphism.

Finally, suppose $Y \subseteq X$ is transitive. Then for $x \in Y$ we have

$$y \in x \rightarrow y \in Y,$$

so for $x \in Y$ we can write the definition of π as

$$\pi(x) = \{\pi(y) \mid y \in x\}.$$

Suppose now that $\pi \upharpoonright Y \neq \text{id} \upharpoonright Y$. Pick $x \in Y$ of minimal rank such that $\pi(x) \neq x$. Then for $y \in x$, we must have $\pi(y) = y$, so

$$\pi(x) = \{\pi(y) \mid y \in x\} = \{y \mid y \in x\} = x,$$

a contradiction. The proof is complete. \square

By an analogous argument we may also prove the following result:

7.2 Theorem (The Representation Lemma). *Let X be a set (or, more generally, a class), E a binary relation on X such that:*

- (i) $(\forall u, v \in X)(u \neq v \rightarrow (\exists x \in X)(xEu \leftrightarrow \neg xEv))$;
- (ii) E is well-founded.

Then there is a unique transitive set (in the general case a class) M and a unique map π such that

$$\pi: \langle X, E \rangle \cong \langle M, \in \rangle.$$

(M is called the *transitivisation* of $\langle X, E \rangle$.) \square

8. Metamathematics of Set Theory

In this section we establish various metamathematical results about the theory ZF. We commence with the well-known *reflection principle* of Montague and Lévy. Loosely speaking, this says that any valid sentence Φ of LST “reflects down” to some initial section V_α of V (i.e. is valid in V_α).

We first of all prove a “generalised reflection principle”. We need some preliminaries.

Let M be any class. For each formula Φ of LST we define a new formula Φ^M , called the *relativisation* of Φ to M . Φ^M is also a formula of LST. This may not be immediately clear from our definition, since it will appear that we are using a unary predicate letter M . But since M is a class, it must be defined by some formula of LST, and by replacing all mention of “ M ” in our definition by this formula we obtain a formula of LST.

The idea is that Φ^M should make the same assertion as Φ , but referring only to the sets in M . In particular, all quantifiers in Φ^M should range only over M . The formal definition of Φ^M proceeds by unravelling the logical construction of Φ and using the following rules.

If Φ is primitive, then $\Phi^M = \Phi$.

If Φ is of the form $\Phi_1 \wedge \Phi_2$, then $\Phi^M = \Phi_1^M \wedge \Phi_2^M$.

If Φ is of the form $\neg \Phi_0$, then $\Phi^M = \neg (\Phi_0^M)$.

If Φ is of the form $\exists v_n \Phi_0$, then $\Phi^M = (\exists v_n \in M)(\Phi_0^M)$. More precisely, if Θ is the LST formula such that

$$M = \{x \mid \Theta(x)\},$$

then Φ^M is the formula

$$\exists v_n [\Theta(v_n) \wedge \Phi_0^M].$$

(This is the really significant clause of course, being the only one which involves M .)

8.1 Theorem (Generalised Reflection Principle). *Let $(W_\alpha \mid \alpha \in \text{On})$ be a hierarchy of transitive sets, definable by a formula, Ψ , of LST in the sense that*

$$W_\alpha = \{x \mid \Psi(x, \alpha)\},$$

and suppose that:

- (i) $\alpha < \beta \rightarrow W_\alpha \subseteq W_\beta$;
- (ii) $\lim(\delta) \rightarrow W_\delta = \bigcup_{\alpha < \delta} W_\alpha$.

Let

$$W = \bigcup_{\alpha \in \text{On}} W_\alpha.$$

Let $\Phi(\vec{v})$ be an LST-formula with free variables amongst \vec{v} . Then the following sentence is a theorem of ZF:

$$(\forall \alpha)(\exists \beta > \alpha)[\text{lim}(\beta) \wedge (\forall \vec{v} \in W_\beta)[\Phi^W(\vec{v}) \leftrightarrow \Phi^{W_\beta}(\vec{v})]].$$

Proof. Let $\Phi(\vec{v})$ be a given formula of LST. Let ${}^3\Phi_0(\vec{x}_0), \dots, \Phi_n(\vec{x}_n)$ be a sequence of LST-formulas such that $\Phi_n = \Phi$ and for each $i = 0, \dots, n$, either Φ_i is a primitive formula or else is obtained from previous formulas in the sequence by a direct application of negation, conjunction, or existential quantification. (The existence of such a sequence follows from the definition of a formula of LST.) We define ordinal-valued functions $f_i(\vec{x}_i)$, $i = 0, \dots, n$, as follows. If Φ_i is primitive or of the form $\neg \Phi_j$ for some $j < i$, or of the form $\Phi_j \wedge \Phi_k$ for some $j, k < i$, let $f_i(\vec{x}_i) = 0$. If $\Phi_i(\vec{x}_i)$ is of the form $\exists y \Phi_j(y, \vec{x}_i)$ for some $j < i$, let $f_i(\vec{x}_i)$ be the least ordinal γ such that

$$(\exists y \in W) \Phi_j^W(y, \vec{x}_i) \rightarrow (\exists y \in W_\gamma) \Phi_j^W(y, \vec{x}_i).$$

Given α now, let $\beta > \alpha$ be a limit ordinal such that for each $i = 0, \dots, n$,

$$(\forall \vec{x}_i \in W_\beta)(f_i(\vec{x}_i) < \beta).$$

Using the Axiom of Collection, it is easy to show that such a β exists. We prove by induction on $i = 0, \dots, n$ that for $\vec{x}_i \in W_\beta$,

$$\Phi_i^W(\vec{x}_i) \leftrightarrow \Phi_i^{W_\beta}(\vec{x}_i).$$

If Φ_i is primitive (in particular, if $i = 0$) this is immediate. And if Φ_i is $\neg \Phi_j$ or $\Phi_j \wedge \Phi_k$ (where $j, k < i$), the induction step is trivial. So suppose that $\Phi_i(\vec{x}_i)$ is $\exists y \Phi_j(y, \vec{x}_i)$, where $j < i$. Let $\vec{x}_i \in W_\beta$.

Assume first that $\Phi_i^W(\vec{x}_i)$. Thus

$$(\exists y \in W) \Phi_j^W(y, \vec{x}_i).$$

3 For the convention regarding expressions such as $\vec{x}_0, \dots, \vec{x}_n$, see page 4.

Since $f_i(\tilde{x}_i) < \beta$, it follows that

$$(\exists y \in W_\beta) \Phi_j^W(y, \tilde{x}_i).$$

But for $y, x_i \in W_\beta$, the induction hypothesis gives

$$\Phi_j^W(y, \tilde{x}_i) \leftrightarrow \Phi_j^{W_\beta}(y, \tilde{x}_i).$$

Hence $\Phi_i^{W_\beta}(\tilde{x}_i)$.

Now assume that $\Phi_i^{W_\beta}(\tilde{x}_i)$. Thus

$$(\exists y \in W_\beta) \Phi_j^{W_\beta}(y, \tilde{x}_i).$$

Using the induction hypothesis, it follows that

$$(\exists y \in W_\beta) \Phi_j^W(y, \tilde{x}_i).$$

But $W_\beta \subseteq W$. Hence

$$(\exists y \in W) \Phi_j^W(y, \tilde{x}_i).$$

In other words, $\Phi_i^W(\tilde{x}_i)$.

The proof is complete. \square

8.2 Corollary (The Reflection Principle). *Let $\Phi(\tilde{v})$ be any formula of LST with free variables amongst \tilde{v} . Then the following sentence is provable in ZF:*

$$(\forall \alpha)(\exists \beta > \alpha) [\lim(\beta) \wedge (\forall \tilde{v} \in V_\beta) [\Phi(\tilde{v}) \leftrightarrow \Phi^{V_\beta}(\tilde{v})]]. \quad \square$$

Using 8.2, it is quite easy to show that ZF is not finitely axiomatisable. For suppose there were a finite set $\{\Phi_1, \dots, \Phi_n\}$ of LST-sentences which yielded all of the ZF axioms. Let

$$\Phi = \Phi_1 \wedge \dots \wedge \Phi_n.$$

Thus Φ is a single axiom for the theory ZF. By the Reflection Principle there is an ordinal α such that Φ^{V_α} . Let α be the least such. Then V_α is a model of ZF. Hence we can apply the Reflection Principle *within* V_α to find an ordinal $\beta \in V_\alpha$ such that, *within* V_α , Φ^{V_β} . But this implies that Φ^{V_β} is valid in the real world, and since $\beta < \alpha$, this contradicts the choice of α , and we are done. Notice that we have made various assumptions in this argument. Firstly, that the set “ V_β ” as constructed within the “universe” V_α is the same as the real V_β , constructed in V . Secondly, we assumed that if Φ^{V_β} is true inside V_α , then Φ^{V_β} really is true. These are easily verified examples of the important general concept of *absoluteness*, discussed below.

Let M be a transitive class and let $\Phi(\tilde{x})$ be an LST-formula. We say $\Phi(\tilde{x})$ is *downward absolute* (*D-absolute*) for M iff

$$(\forall \tilde{x} \in M) (\Phi(\tilde{x}) \rightarrow \Phi^M(\tilde{x})).$$

We say $\Phi(\vec{x})$ is *upward absolute (U-absolute) for M* iff

$$(\forall \vec{x} \in M)(\Phi^M(\vec{x}) \rightarrow \Phi(\vec{x})).$$

Finally, we say $\Phi(\vec{x})$ is *absolute for M* iff it is both *D-absolute* and *U-absolute* for *M*.

In cases where it is clear which class *M* is concerned, we often drop the explicit mention of *M*, and say, for example, simply that “ $\Phi(\vec{x})$ is absolute”.

It is clear that primitive LST-formulas will be absolute for any class *M*. For other formulas we usually need to assume that the class *M* is transitive; in which case absoluteness is related to the logical complexity of the formula. In this connection, a classification of the logical complexity of formulas due to Lévy is useful.

In order to describe the *Lévy hierarchy* in the simplest fashion, it is convenient to regard both universal quantifiers and the two types of bounded quantifier as integral parts of LST, rather than as mere abbreviations.

Let Φ be an LST-formula. We say that Φ is Σ_0 (or Π_0) iff it contains no unbounded quantifiers. Thus the only quantifiers in Φ will be of the form $(\forall x \in y)$ or $(\exists x \in y)$. Now let $n \geq 1$. Recursively, we say that Φ is Σ_n iff it is of the form $\exists \vec{x} \Psi(\vec{x})$ where $\Psi(\vec{x})$ is Π_{n-1} , and that Φ is Π_n iff it is of the form $\forall \vec{x} \Psi(\vec{x})$ where $\Psi(\vec{x})$ is Σ_{n-1} .

Thus, to say that a formula is Σ_n is to say that the formula consists of a Σ_0 formula preceded by *n* blocks of like quantifiers, commencing with a block of existential quantifiers and alternating between blocks of existential quantifiers and blocks of universal quantifiers.

A formula Φ is said to be Σ_n^{ZF} iff there is a Σ_n formula Ψ such that

$$\text{ZF} \vdash \Phi \leftrightarrow \Psi.$$

Similarly Π_n^{ZF} . For $n \geq 1$, a formula Φ is said to be Δ_n^{ZF} iff it is both Σ_n^{ZF} and Π_n^{ZF} .

If *T* is some subtheory of ZF, we define $\Sigma_n^T, \Pi_n^T, \Delta_n^T$ in a similar fashion, with *T* in place of ZF.

The following simple result is fundamental to much of our later work.

8.3 Lemma. *Let T be some subtheory of ZF (possibly ZF itself). Let M be a transitive class such that Ψ^M for every axiom Ψ of T. Let Φ be any formula of LST.*

- (i) *If Φ is Σ_0^T , then Φ is absolute (for M).*
- (ii) *If Φ is Σ_1^T , then Φ is U-absolute.*
- (iii) *If Φ is Π_1^T , then Φ is D-absolute.*
- (iv) *If Φ is Δ_1^T , then Φ is absolute.*

Proof. (i) Let Φ have its free variables amongst \vec{v} , and let $\Psi(\vec{v})$ be a Σ_0 formula of LST such that

$$T \vdash \Phi \leftrightarrow \Psi.$$

Since *T* is a subtheory of ZF,

$$\forall \vec{v} [\Phi \leftrightarrow \Psi]$$

is a valid assertion. And since Θ^M for every axiom Θ of T ,

$$(\forall \bar{v} [\Phi \leftrightarrow \Psi])^M$$

is a valid assertion, which is to say that

$$(\forall \bar{v} \in M) [\Phi^M \leftrightarrow \Psi^M]$$

is valid. Hence it suffices to prove the result for Ψ . In other words, there is no loss of generality in assuming that Φ is itself a Σ_0 formula.

If Φ is primitive, the result is immediate. We proceed now by induction on the construction of Φ . If Φ is of the form $\Phi_1 \wedge \Phi_2$ or of the form $\neg \Phi_0$, the result for Φ follows trivially from the result for Φ_0, Φ_1, Φ_2 . Suppose that $\Phi(y, \bar{v})$ has the form $(\exists x \in y) \Psi(x, y, \bar{v})$, where the result is valid for Ψ . Let $y, \bar{v} \in M$. If $\Phi(y, \bar{v})^M$, then $[(\exists x \in y) \Psi(x, y, \bar{v})]^M$ so for some $x \in M$ such that $x \in y$ we have $\Psi(x, y, \bar{v})^M$. So by induction hypothesis, $\Psi(x, y, \bar{v})$. Hence $(\exists x \in y) \Psi(x, y, \bar{v})$, which means that $\Phi(y, \bar{v})$. Conversely, suppose $\Phi(y, \bar{v})$. Thus $(\exists x \in y) \Psi(x, y, \bar{v})$, so for some $x \in y$, $\Psi(x, y, \bar{v})$. But M is transitive, so as $y \in M$ we have $x \in M$ also. (Note the importance of transitivity here.) Hence by the induction hypothesis we can conclude that $\Psi(x, y, \bar{v})^M$, and hence that $[(\exists x \in y) \Psi(x, y, \bar{v})]^M$, i.e. $\Phi(y, \bar{v})^M$. The case where Φ has the form $(\forall x \in y) \Psi(x, y, \bar{v})$ is handled similarly. This proves (i).

(ii) As in part (i) there is no loss in generality in assuming that Φ is a Σ_1 formula. Let $\Phi = \Phi(\bar{v}) = \exists \bar{u} \Psi(\bar{u}, \bar{v})$, where Ψ is Σ_0 . Assume $\Phi(\bar{v})^M$, where $\bar{v} \in M$. Thus for some $\bar{u} \in M$, $\Psi(\bar{u}, \bar{v})^M$. By part (i), it follows that $\Psi(\bar{u}, \bar{v})$. Thus $\exists \bar{u} \Psi(\bar{u}, \bar{v})$, i.e. $\Phi(\bar{v})$.

(iii) As before we may assume that Φ is Π_1 . Let $\Phi = \Phi(\bar{v}) = \forall \bar{u} \Psi(\bar{u}, \bar{v})$, where Ψ is Σ_0 . Assume $\Phi(\bar{v})$, where $\bar{v} \in M$. Then for all \bar{u} , $\Psi(\bar{u}, \bar{v})$. In particular, for all $\bar{u} \in M$, $\Psi(\bar{u}, \bar{v})$. But by part (i), if $\bar{u} \in M$, then $\Psi(\bar{u}, \bar{v})$ implies $\Psi(\bar{u}, \bar{v})^M$. Hence for all $\bar{u} \in M$, $\Psi(\bar{u}, \bar{v})^M$. In other words, $\Phi(\bar{v})^M$.

(iv) By parts (ii) and (iii). \square

Let us see where many of the simpler formulas of LST lie in the Lévy hierarchy. Notice that when we speak of, for example

the formula “ x is a finite sequence”,

we mean the “obvious” rendering of this statement as a formula in LST. In most cases the rendering is indeed *obvious*. If there is any significant doubt, we shall indicate the manner in which the statement is expressed in LST. In the case of our first lemma there is no such problem.

8.4. Lemma. *The following formulas are Σ_0 : $x = y$, $x \in y$, $x \subseteq y$, $y = \{x_1, \dots, x_n\}$, $y = (x_1, \dots, x_n)$, $y = (x)_i^n$ (for $i = 0, \dots, n-1$), $y = x \cup z$, $y = x \cap z$, $y = \bigcup x$, $y = \bigcap x$, $y = x - z$, “ x is an n -tuple”, “ x is a relation on y ”, “ x is a function”, $y = \text{dom}(x)$, $y = \text{ran}(x)$, $y = x(z)$, $y = x''z$, $y = x \upharpoonright z$, $y = x \times z$, $y = x^{-1}$, $y = x \cup \{x\}$, $\text{On}(x)$, $\text{lim}(x)$, $\text{succ}(x)$, “ x is a natural number”, “ x is a sequence”, $x : y \rightarrow z$, $x : y \leftrightarrow z$. \square*

In the case of the next lemma, we shall indicate the fashion in which the statement is to be expressed in LST, since, unlike the previous lemma, there are several possibilities.

8.5 Lemma. *The formula “ x is finite” is Σ_1 .*

Proof. x is finite $\leftrightarrow \exists n \exists f [n \text{ is a natural number} \wedge f: n \leftrightarrow x]$. \square

The following lemma gives various closure properties for the levels in the Lévy hierarchy. The proofs are all trivial.

8.6 Lemma. *Let T be any LST theory. (By convention, T therefore includes all the axioms for predicate logic for LST.) Let Φ, Ψ be formulas of LST.*

- (i) *If Φ, Ψ are Σ_0^T , so too are $\Phi \wedge \Psi, \Phi \vee \Psi, \neg \Phi$.*
- (ii) *If Φ is $\Sigma_n^T, \neg \Phi$ is Π_n^T ; if Φ is $\Pi_n^T, \neg \Phi$ is Σ_n^T .*
- (iii) *Φ is Δ_n^T iff both Φ and $\neg \Phi$ are Σ_n^T .*
- (iv) *If Φ, Ψ are Σ_n^T , so are $\Phi \wedge \Psi, \Phi \vee \Psi, \exists x \Phi, (\exists x \in z) \Phi$.*
- (v) *If Φ, Ψ are Π_n^T , so are $\Phi \wedge \Psi, \Phi \vee \Psi, \forall x \Phi, (\forall x \in z) \Phi$.*
- (vi) *If Φ, Ψ are Δ_n^T , so are $\Phi \wedge \Psi, \Phi \vee \Psi, \neg \Phi$.*
- (vii) *$m < n \rightarrow \Sigma_m^T, \Pi_m^T \subseteq \Delta_n^T$. \square*

8.7 Lemma. *The formula*

$$\text{WF}(x, y) \leftrightarrow \text{“}x \text{ is a well-founded relation on } y\text{”}$$

is Δ_1^{ZF} .

Proof. It is easily seen that the formula

$$\text{“}x \text{ is a binary relation on } y\text{”}$$

is Σ_0 . Consequently, we need only concentrate upon the clause of WF which relates to well-foundedness. Let Φ denote this clause. Now, if E is a binary relation on a set X , the obvious rendering of $\Phi(E, X)$ is:

$$\forall A [A \subseteq X \wedge A \neq \emptyset \rightarrow (\exists a \in A)(\forall x \in A) \neg (x E a)].$$

(This is the *definition* of well-foundedness.) This shows at once that in its canonical rendering in LST, the formula $\text{WF}(x, y)$ is Π_1 . But in ZF it is easy to prove, for E, X as above, the equivalence

$$\Phi(E, X) \leftrightarrow \exists f [f: X \rightarrow \text{On} \wedge (\forall x, y \in X)(x E y \rightarrow f(x) < f(y))].$$

(This involves a fairly routine application of the Recursion Principle.) This shows that Φ , and hence WF, are Σ_1^{ZF} , so we are done. \square

Given an LST formula $\Phi(y, \vec{z})$, we denote by $\Phi((x)_0, \vec{z})$ the LST-formula

$$(\exists u \in x)(\exists a \in u)(\exists b \in u)[x = (a, b) \wedge \Phi(a, \vec{z})].$$

Similarly for $\Phi((x)_1, \vec{z})$.

Again, we denote by $\Phi(x(y), \vec{z})$ the LST-formula

$$(x \text{ is a function}) \wedge (\exists w \in x)(\exists u \in w)(\exists v \in u)[w = (v, y) \wedge \Phi(v, \vec{z})].$$

The following lemma is an immediate consequence of these definitions.

8.8 Lemma. *If $\Phi(x, \vec{z})$ is a Σ_0 formula of LST, then so too are $\Phi((x)_0, \vec{z})$, $\Phi((x)_1, \vec{z})$, and $\Phi(x(y), \vec{z})$. \square*

8.9 Lemma (Contraction of Quantifiers). *Let T be any LST theory whose axioms include the axioms of null set and pairing (see sections 2 and 3). Then:*

(i) *Let $n \geq 1$ and let $\Phi(\vec{z})$ be a Σ_n formula. Then there is a Σ_0 formula $\Psi(\vec{x}, \vec{z})$ such that*

$$T \vdash \Phi(\vec{z}) \leftrightarrow \exists x_1 \forall x_2 \exists x_3 \dots - x_n \Psi(x_1, \dots, x_n, \vec{z}).$$

(ii) *Let $n \geq 1$ and let $\Phi(\vec{z})$ be a Π_n formula. Then there is a Σ_0 formula $\Phi(\vec{x}, \vec{z})$ such that*

$$T \vdash \Phi(\vec{z}) \leftrightarrow \forall x_1 \exists x_2 \forall x_3 \dots - x_n \Psi(x_1, \dots, x_n, \vec{z}).$$

Proof. We prove (i). The proof of (ii) is similar. Consider first the case $n = 1$. A general Σ_1 formula has the form

$$\Phi(\vec{z}): \exists y_1 \exists y_2 \dots \exists y_m \Theta(y_1, \dots, y_m, \vec{z}),$$

where Θ is Σ_0 . If $m = 1$ now there is nothing further to prove. Suppose that $m = 2$. (All other cases $m > 2$ are handled similarly.) Let $\Psi(x, \vec{z})$ be the formula:

$$(x \text{ is an ordered pair}) \wedge \Theta((x)_0, (x)_1, \vec{z}).$$

By 8.8, Ψ is Σ_0 . And clearly, by our assumptions on T ,

$$T \vdash \Phi(\vec{z}) \leftrightarrow \exists x \Psi(x, \vec{z}).$$

That deals with the case $n = 1$. We consider next the case $n = 2$, and leave it to the reader to see that the same idea works for all cases $n \geq 2$. Suppose $\Phi(\vec{z})$ is the formula

$$\exists u_1 \exists u_2 \dots \exists u_p \forall v_1 \forall v_2 \dots \forall v_q \Theta(\vec{u}, \vec{v}, \vec{z}),$$

where Θ is Σ_0 . Let $\Psi(x, y, \vec{z})$ be the formula

$$(x \text{ is a } p\text{-tuple}) \wedge [(y \text{ is a } q\text{-tuple}) \\ \rightarrow \Theta((x)_0^p, \dots, (x)_{p-1}^p, (y)_0^q, \dots, (y)_{q-1}^q, \vec{z})].$$

By 8.8, Ψ is Σ_0 . Moreover,

$$T \vdash \Phi(\vec{z}) \leftrightarrow \exists x \forall y \Psi(x, y, \vec{z}).$$

The proof is complete. \square

For the case where the theory, T , concerned is ZF, the following lemma extends the closure rules given in 8.6.

8.10 Lemma.

- (i) If Φ is a Σ_n formula of LST, then $(\forall x \in y) \Phi$ is Σ_n^{ZF} .
(ii) If Φ is a Π_n formula of LST, then $(\exists x \in y) \Phi$ is Π_n^{ZF} .

Proof. We prove (i) and (ii) simultaneously by induction on n . For $n = 0$ there is nothing to prove. Suppose now that (i) and (ii) hold for n . We prove (i) and (ii) for $n + 1$.

- (i) Let Φ be Σ_{n+1} . By 8.9 there is a Π_n formula Ψ such that

$$\text{ZF} \vdash \Phi \leftrightarrow \exists z \Psi.$$

Hence,

$$\text{ZF} \vdash (\forall x \in y) \Phi \leftrightarrow (\forall x \in y) \exists z \Psi.$$

But, by using the Axiom of Collection,

$$\text{ZF} \vdash (\forall x \in y) \exists z \Psi \leftrightarrow \exists u (\forall x \in y) (\exists z \in u) \Psi.$$

Thus

$$\text{ZF} \vdash (\forall x \in y) \Phi \leftrightarrow \exists u (\forall x \in y) (\exists z \in u) \Psi.$$

By induction hypothesis, $(\exists z \in u) \Psi$ is Π_n^{ZF} . Hence, using 8.7, $(\forall x \in y) (\exists z \in u) \Psi$ is Π_n^{ZF} . Thus $\exists u (\forall x \in y) (\exists z \in u) \Psi$ is Σ_{n+1}^{ZF} , which means that $(\forall x \in y) \Phi$ is Σ_{n+1}^{ZF} , as required.

(ii) Now suppose that Φ is Π_{n+1} . Then $\neg \Phi$ is Σ_{n+1}^{ZF} . Hence by the above, $(\forall x \in y) \neg \Phi$ is Σ_{n+1}^{ZF} . It follows that $\neg (\exists x \in y) \Phi$ is Σ_{n+1}^{ZF} , and hence that $(\exists x \in y) \Phi$ is Π_{n+1}^{ZF} . \square

9. The Language \mathcal{L}_V

We develop, within set theory, a formal “language”, \mathcal{L}_V , which consists of an analogue of LST (which analogue we shall denote by \mathcal{L}), together with an individual constant “symbol” for each set (in V). The purpose of the subscript V in “ \mathcal{L}_V ” is to indicate that these constants are present. Later on we shall consider sub-languages \mathcal{L}_X of \mathcal{L}_V for any set X , where we only allow constants which denote elements of X . In particular, \mathcal{L}_\emptyset is the same as \mathcal{L} , the formal analogue of LST.

It should be emphasised that the entire development of \mathcal{L}_V takes place within set theory. In particular, all the “symbols” and “formulas” of \mathcal{L}_V will be sets. We shall require that the various syntactic and semantic notions of \mathcal{L}_V have certain absoluteness properties, and in order to see that this is the case we shall need to examine the logical complexity of the (real) LST formulas which define the various

notions of \mathcal{L}_V . So, as we proceed with our development of \mathcal{L}_V within set theory, we shall make regular metamathematical digressions to examine the logical structure of the various notions. To try to minimise any confusion, we shall use lower case Greek letters $\varphi, \psi, \theta, \dots$ to denote “formulas” of \mathcal{L}_V (with upper case Greek letters $\Phi, \Psi, \Theta, \dots$ for formulas of LST as before). However, since \mathcal{L} will have the same structure as LST, it would be an unnecessary complication to use separate symbols for the variables and connectives of these two languages, so we shall leave this distinction to the reader, who will always be aided by the context.

The basic symbols of \mathcal{L}_V will be certain sets, and the formulas of \mathcal{L}_V will be certain finite sequences of these sets. Accordingly, we must begin by establishing some notations concerning finite sequence. (Incidentally, the exact fashion in which \mathcal{L}_V is defined is not important, and we have just chosen a reasonably convenient method.)

The sequence with domain $\{0\}$ and value x is denoted by $\langle x \rangle$. The finite sequence with domain $\{0, \dots, n-1\}$ and values x_0, \dots, x_{n-1} is denoted by $\langle x_0, \dots, x_{n-1} \rangle$. (Notice that $\langle x_0, \dots, x_{n-1} \rangle$ is not the same as the n -tuple (x_0, \dots, x_{n-1}) .)

If s, t are sequences, $s \frown t$ denotes the *concatenation* of s and t , i.e. if $s = \langle x_0, \dots, x_{n-1} \rangle$ and $t = \langle y_0, \dots, y_{m-1} \rangle$, then

$$s \frown t = \langle x_0, \dots, x_{n-1}, y_0, \dots, y_{m-1} \rangle.$$

If s is a finite sequence, $\|s\|$ denotes the greatest element of $\text{dom}(s)$, i.e. $\|s\| = \text{dom}(s) - 1$.

The *variables* of the language \mathcal{L}_V are the sets $(2, n)$, for $n \in \omega$, and we shall denote $(2, n)$ by the symbol v_n .

Let $\text{Vbl}(x)$ be the following LST formula:

$$[x \text{ is an ordered pair}] \wedge [(x)_0 = 2] \wedge [(x)_1 \text{ is a natural number}].$$

Clearly,

$$\text{Vbl}(x) \leftrightarrow x \text{ is a variable of } \mathcal{L}_V.$$

In the above equivalence, note the use of the symbol x . Vbl is an LST formula, and x is a variable of LST. Being a variable of LST, x denotes a set. $\text{Vbl}(x)$ says that the set denoted by x has the form that we have decided to refer to as a “variable” of \mathcal{L}_V . With a little experience, any initial difficulties the reader may encounter due to points such as this should be easily overcome.

For each set x , \mathcal{L}_V has an *individual constant symbol*, namely the set $(3, x)$, which we shall denote by \dot{x} .

Let $\text{Const}(x)$ be the LST formula:

$$[x \text{ is an ordered pair}] \wedge [(x)_0 = 3].$$

Clearly,

$$\text{Const}(x) \leftrightarrow x \text{ is a constant of } \mathcal{L}_V.$$

The *primitive formulas* of \mathcal{L}_V are the sequences of the forms

$$\langle 0, 4, x, y, 1 \rangle \quad \text{and} \quad \langle 0, 5, x, y, 1 \rangle,$$

where x and y are variables or constants of \mathcal{L}_V . The sequence $\langle 0, 4, x, y, 1 \rangle$ will be denoted by $(x \in y)$, and the sequence $\langle 0, 5, x, y, 1 \rangle$ by $(x = y)$. (Thus we are using the number 0 to correspond to the open bracket symbol of LST and the number 1 to correspond to the close bracket symbol. The number 4 indicates a membership formula, and the number 5 indicates an equality formula.)

Let $\text{PFml}(x)$ be the LST formula:

$$\begin{aligned} & [x \text{ is a function}] \wedge [\text{dom}(x) = 5] \wedge [x(0) = 0] \wedge [x(1) = 4 \vee x(1) = 5] \\ & \wedge [\text{Vbl}(x(2)) \vee \text{Const}(x(2))] \wedge [\text{Vbl}(x(3)) \vee \text{Const}(x(3))] \\ & \wedge [x(4) = 1]. \end{aligned}$$

Clearly,

$$\text{PFml}(x) \leftrightarrow x \text{ is a primitive formula of } \mathcal{L}_V.$$

9.1 Lemma. *The LST formulas $\text{Vbl}(x)$, $\text{Const}(x)$, $\text{PFml}(x)$ are all Σ_0 (when written out fully in LST).*

Proof. Immediate. \square

The *formulas* of \mathcal{L}_V are built up from the primitive formulas by means of the following schemas:

$$\begin{aligned} (\varphi \wedge \psi) &= \langle 0, 6 \rangle \frown \varphi \frown \psi \frown \langle 1 \rangle \\ (\neg \varphi) &= \langle 0, 7 \rangle \frown \varphi \frown \langle 1 \rangle \\ (\exists u \varphi) &= \langle 0, 8, u \rangle \frown \varphi \frown \langle 1 \rangle, \end{aligned}$$

where φ, ψ are formulas of \mathcal{L}_V and u is a variable of \mathcal{L}_V . (Note again the use of 0 and 1 as brackets, with the numbers 6, 7, 8 indicating the operations of conjunction, negation, and existential quantification, respectively.)

We shall presently write down a Σ_1 formula of LST which says “ x is a formula of \mathcal{L}_V ”. But before we can do this we require several preliminary notions.

The following LST formula, $\text{Finseq}(x)$, says that “ x is a finite sequence”:

$$\begin{aligned} & [x \text{ is a sequence}] \wedge (\forall u \in \text{dom}(x)) [u \text{ is a natural number}] \\ & \wedge (\exists u \in \text{dom}(x)) (\forall u \in \text{dom}(x)) [u \in v \vee u = v]. \end{aligned}$$

9.2 Lemma. *The LST formula $\text{Finseq}(x)$ is (when written out fully in LST) Σ_0 .*

Proof. All we need to observe is that expressions such as

$$(\forall u \in \text{dom}(x)) [\dots u \dots]$$

can be replaced by

$$(\forall z \in x) [\dots (z)_1 \dots],$$

which is Σ_0 by 8.8. \square

We now write down LST formulas which describe the construction of the “formulas” of \mathcal{L}_V .

Let $F_\epsilon(\theta, x, y)$ be the LST formula

$$\text{Finseq}(\theta) \wedge [\text{dom}(\theta) = 5] \wedge [\theta(0) = 0] \wedge [\theta(1) = 4] \wedge [\theta(2) = x] \\ \wedge [\theta(3) = y] \wedge [\theta(4) = 1].$$

Clearly, if $x, y \in \text{Vbl} \cup \text{Const}$, then

$$F_\epsilon(\theta, x, y) \leftrightarrow \theta \text{ is the } \mathcal{L}_V \text{ formula } (x \in y).$$

Let $F_=(\theta, x, y)$ be the LST formula

$$\text{Finseq}(\theta) \wedge [\text{dom}(\theta) = 5] \wedge [\theta(0) = 0] \wedge [\theta(1) = 5] \wedge [\theta(2) = x] \\ \wedge [\theta(3) = y] \wedge [\theta(4) = 1].$$

Thus if $x, y \in \text{Vbl} \cup \text{Const}$, then

$$F_=(\theta, x, y) \leftrightarrow \theta \text{ is the } \mathcal{L}_V \text{ formula } (x = y).$$

Let $F_\wedge(\theta, \varphi, \psi)$ be the LST formula

$$\text{Finseq}(\theta) \wedge \text{Finseq}(\varphi) \wedge \text{Finseq}(\psi) \\ \wedge [\text{dom}(\theta) = \text{dom}(\varphi) + \text{dom}(\psi) + 3] \wedge [\theta(0) = 0] \wedge [\theta(1) = 6] \\ \wedge [\theta(\|\theta\|) = 1] \wedge (\forall i \in \text{dom}(\varphi))[\theta(i + 2) = \varphi(i)] \\ \wedge (\forall i \in \text{dom}(\psi))[\theta(\text{dom}(\varphi) + i + 2) = \psi(i)].$$

Thus if $\varphi, \psi \in \text{Fml}$, then

$$F_\wedge(\theta, \varphi, \psi) \leftrightarrow \theta \text{ is the } \mathcal{L}_V \text{ formula } (\varphi \wedge \psi).$$

Let $F_\neg(\theta, \varphi)$ be the LST formula

$$\text{Finseq}(\theta) \wedge \text{Finseq}(\varphi) \wedge [\text{dom}(\theta) = \text{dom}(\varphi) + 3] \wedge [\theta(0) = 0] \\ \wedge [\theta(1) = 7] \wedge [\theta(\|\theta\|) = 1] \wedge (\forall i \in \text{dom}(\varphi))[\theta(i + 2) = \varphi(i)].$$

Thus if $\varphi \in \text{Fml}$, then

$$F_\neg(\theta, \varphi) \leftrightarrow \theta \text{ is the } \mathcal{L}_V \text{ formula } (\neg \varphi).$$

Finally, let $F_\exists(\theta, u, \varphi)$ be the LST formula

$$\text{Finseq}(\theta) \wedge \text{Finseq}(\varphi) \wedge [\text{dom}(\theta) = \text{dom}(\varphi) + 4] \wedge [\theta(0) = 0] \\ \wedge [\theta(1) = 8] \wedge [\theta(2) = u] \wedge [\theta(\|\theta\|) = 1] \\ \wedge (\forall i \in \text{dom}(\varphi))[\theta(i + 3) = \varphi(i)].$$

Thus if $\varphi \in \text{Fml}$ and $u \in \text{Vbl}$, we have

$$F_{\exists}(\theta, u, \varphi) \leftrightarrow \theta \text{ is the } \mathcal{L}_V \text{ formula } (\exists u \varphi).$$

9.3 Lemma. *The LST formulas $F_{=}$, F_{\neq} , F_{\wedge} , F_{\neg} , F_{\exists} are all Σ_0 (when written out fully in LST).*

Proof. In view of the remark made in the proof of 9.2, this is clear from the nature of the formulas concerned. \square

Now, if φ is a formula of \mathcal{L}_V , there must be a finite sequence ψ_0, \dots, ψ_n of \mathcal{L}_V formulas such that $\psi_n = \varphi$ and for each i , ψ_i is either a primitive formula or else is obtained from one or two formulas in the list $\psi_0, \dots, \psi_{i-1}$ by an application of one of the schemas for generating formulas. The sequence ψ_0, \dots, ψ_n thus describes the way that φ is built up as a formula. We write down an LST formula, $\text{Build}(\varphi, \psi)$ which says that ψ is just such a sequence ψ_0, \dots, ψ_n . $\text{Build}(\varphi, \psi)$ is as follows:

$$\begin{aligned} \text{Finseq}(\psi) \wedge [\psi_{\parallel\psi\parallel} = \varphi] \wedge (\forall i \in \text{dom}(\psi)) [& \text{PFml}(\psi_i) \\ & \vee (\exists j, k \in i) F_{\wedge}(\psi_i, \psi_j, \psi_k) \vee (\exists j \in i) F_{\neg}(\psi_i, \psi_j) \\ & \vee (\exists j \in i) (\exists u \in \text{ran}(\varphi)) (\text{Vbl}(u) \wedge F_{\exists}(\psi_i, u, \psi_j))]. \end{aligned}$$

9.4 Lemma. *The LST formula $\text{Build}(\varphi, \psi)$ is Σ_0 (when written out fully in LST).*

Proof. The main point to check is that expressions such as

$$(\forall i \in \text{dom}(\psi)) (\exists j, k \in i) F_{\wedge}(\psi_i, \psi_j, \psi_k)$$

are Σ_0 . Well, this one can be written as

$$\begin{aligned} (\forall i \in \text{dom}(\psi)) (\exists j, k \in i) (\exists a, b, c \in \text{ran}(\psi)) [& a = \psi_j \wedge b = \psi_k \wedge c \\ & \wedge c = \psi_i \wedge F_{\wedge}(a, b, c)], \end{aligned}$$

which is immediately recognisable as Σ_0 now. The other cases are handled similarly. \square

Clearly,

$$\varphi \text{ is a formula of } \mathcal{L}_V \leftrightarrow (\exists f) \text{Build}(\varphi, f),$$

which presents us with a Σ_1 formula of LST to define the formulas of \mathcal{L}_V . Now, our main purpose in analysing the logical complexity of the syntactic notions of \mathcal{L}_V is to enable us to prove various absoluteness results. In the case of Σ_0 notions, such as in Lemmas 9.1 through 9.4, there is no problem, since then 8.3(i) guarantees absoluteness for all transitive classes. But for notions which are not Σ_0 , such as the notion of being a formula of \mathcal{L} , it is not enough to know that the concept is Σ_1 , for that will only guarantee U -absoluteness (see 8.3(ii)). For full absoluteness we require (see 8.3(iv)) an equivalent Π_1 definition as well. Moreover (see 8.3 again), in order that any absoluteness results have the widest possible application,

it is important that the equivalence of the Σ_1 and Π_1 definitions be proved in the simplest theory possible, thereby giving absoluteness for all transitive models of that theory. We now develop such a theory: we call it *Basic Set Theory* (BS).

BS is the LST theory having the following axioms:

- (1) *Extensionality*: $\forall x \forall y [\forall z (z \in x \leftrightarrow z \in y) \rightarrow (x = y)]$;
- (2) *Induction schema*: $\forall \tilde{a} [\forall x ((\forall y \in x) \Phi(y, \tilde{a}) \rightarrow \Phi(x, \tilde{a})) \rightarrow \forall x \Phi(x, \tilde{a})]$, where Φ is any formula of LST with free variables amongst x, \tilde{a} ;
- (3) *Pairing*: $\forall x \forall y \exists z \forall w [(w \in z) \leftrightarrow (w = x \vee w = y)]$;
- (4) *Union*: $\forall x \exists y \forall z [(z \in y) \leftrightarrow (\exists u \in x)(z \in u)]$;
- (5) *Infinity*: $\exists x [\text{On}(x) \wedge (x \neq 0) \wedge (\forall y \in x)(\exists z \in x)(y \in z)]$;
- (6) *Cartesian Product*: $\forall x \forall y \exists z \forall u [(u \in z) \leftrightarrow (\exists a \in x)(\exists b \in y)(u = (a, b))]$;
- (7) Σ_0 -*Comprehension (schema)*: $\forall \tilde{a} \forall x \exists y \forall z [(z \in y) \leftrightarrow (z \in x \wedge \Phi(\tilde{a}, z))]$, where $\Phi(\tilde{a}, z)$ is a Σ_0 formula of LST.

Clearly, BS is a subtheory of ZF (i.e. all the axioms of BS are theorems of ZF). Indeed, axioms (1), (4) and (5) are axioms of ZF, though in the present formulation of the Axiom of Infinity we axiomatically guarantee the existence of an infinite ordinal, rather than any infinite set as we did with ZF. Axioms (3) and (6), which guarantee the existence of the unordered pair $\{x, y\}$ and the Cartesian product $x \times y$ of any two sets x and y , respectively, are easily proved theorems of ZF. Axiom (7) is just the restriction of the usual Comprehension Axiom Schema to the Σ_0 formulas of LST. In the absence of full Comprehension, we replace the Axiom of Foundation of ZF by the induction schema (2).

Notice that BS allows for the construction of all finite sets, i.e. for any n ,

$$\text{BS} \vdash \forall x_1 \dots \forall x_n \exists y \forall z [(z \in y) \leftrightarrow (z = x_1 \vee \dots \vee z = x_n)].$$

We now write down a formula of LST, $\text{Seq}(u, a, n)$, which says that u is the set of all m -sequences of members of a for all $m < n$. Now, the “obvious” formula which says this is:

$$\begin{aligned} & (\forall x \in u)(\exists m \in n)(x \text{ is an } m\text{-sequence of members of } a) \\ & \wedge (\forall x)(\forall m \in n)(x \text{ is an } m\text{-sequence of members of } a \rightarrow x \in u). \end{aligned}$$

But this formula is Π_1 , whereas we shall require a Σ_1 definition. (Though we shall show that our Σ_1 definition is in fact BS-provably equivalent to a Π_1 definition.) We obtain our desired Σ_1 formula by regarding the members of the set u being built up in stages, constructing first the 1-sequences, then the 2-sequences, and so on. (The function f in the following formula enumerates these sets of finite sequences.)

Let $\text{Seq}(u, a, n)$ be the following formula of LST:

$$\begin{aligned} & (\exists f) [\text{Finseq}(f) \wedge (n \text{ is a natural number}) \wedge (\text{dom}(f) = n) \\ & \wedge (u = \bigcup \text{ran}(f)) \wedge (\forall i \in \text{dom}(f))(\forall x \in f(i))(\text{Finseq}(x) \wedge (\text{dom}(x) = i) \\ & \wedge (\forall j \in i)(x(j) \in a)) \wedge (\forall i \in \text{dom}(f))(\forall j \in i)(\forall x \in f(j))(\forall p \in a) \\ & \cdot (i = j + 1 \rightarrow x \cup \{(p, i)\} \in f(i))]. \end{aligned}$$

It is easily seen that this formula is Σ_1 (when written out fully in LST).

9.5 Lemma. *The LST formula $\text{Seq}(u, a, n)$ is Δ_1^{BS} .*

Proof. The only unbounded quantifier in the above formula is $(\exists f)$. This quantifier can, without any loss of generality, be restricted to range over the set of n -sequences of finite sequences from a . (That is to say, if such an f exists, it will have to lie in this bounding set.) Consequently, it is clear from the definition of BS that:

$$\text{BS} \vdash (\forall a)(\forall n \in \omega)(\exists u) \text{Seq}(u, a, n).$$

But we obviously have:

$$\text{BS} \vdash \forall a \forall n \forall u \forall v [\text{Seq}(u, a, n) \wedge \text{Seq}(v, a, n) \rightarrow u = v].$$

Hence,

$$\text{BS} \vdash \text{Seq}(u, a, n) \leftrightarrow [(n \text{ is a natural number}) \wedge \forall z [\text{Seq}(z, a, n) \rightarrow z = u]].$$

This proves the lemma, since the expression on the right of this equivalence is Π_1 . \square

We are now able to write down an LST formula $\text{Fml}(x)$ such that:

$$\text{Fml}(x) \leftrightarrow x \text{ is a formula of } \mathcal{L}_V.$$

As we mentioned earlier, the obvious way to do this is by the formula

$$(\exists f) \text{Build}(x, f).$$

So let us take this as our formula $\text{Fml}(x)$. By 9.4, $\text{Fml}(x)$ is Σ_1 .

9.6 Lemma. *The LST formula $\text{Fml}(x)$ is Δ_1^{BS} .*

Proof. Consider the quantifier $\exists f$ in the expression

$$(\exists f) \text{Build}(x, f).$$

We may clearly bind this quantifier by the set

$$A(x) = \bigcup_{n \in \text{dom}(x)}^{n+1} \left[\bigcup_{n \in \text{dom}(x)}^{m+1} \text{ran}(x) \right].$$

(Because, as is easily seen,

$$(\exists f) \text{Build}(x, f) \rightarrow (\exists f \in A(x)) \text{Build}(x, f).)$$

Moreover, it is easily checked that

$$\text{BS} \vdash \forall x \exists y [y = A(x)].$$

Hence

$$\begin{aligned} \text{BS} \vdash \text{Fml}(x) \leftrightarrow & \text{Finseq}(x) \wedge \forall u \forall v [\text{Seq}(u, \text{ran}(x), \text{dom}(x) + 2) \\ & \wedge \text{Seq}(v, u, \text{dom}(x) + 2) \rightarrow (\exists f \in v) \text{Build}(x, f)]. \end{aligned}$$

This provides us with a Π_1^{BS} equivalent to $\text{Fml}(x)$, so we are done. \square

The above trick of finding a convenient bound for a quantifier will be used frequently during our development of \mathcal{L}_V .

Given any class X , \mathcal{L}_X is the “sublanguage” of \mathcal{L}_V obtained by omitting from \mathcal{L}_V all constants \dot{z} for $z \notin X$. We write \mathcal{L} instead of \mathcal{L}_\emptyset . Thus \mathcal{L} is a formal analogue of LST within set theory. We shall be particularly concerned with the languages \mathcal{L}_u where u is a set.

Let $\text{Const}(x, u)$ be the LST formula

$$\text{Const}(x) \wedge (x)_1 \in u.$$

The LST formulas $\text{PFml}(x, u)$ and $\text{Fml}(x, u)$ are defined in exactly the same way as the formulas $\text{PFml}(x)$ and $\text{Fml}(x)$ except that $\text{Const}(x)$ is replaced everywhere by $\text{Const}(x, u)$. Clearly,

$$\text{Fml}(x, u) \leftrightarrow x \text{ is a formula of } \mathcal{L}_u.$$

By means of arguments as before, we have:

9.7 Lemma.

- (i) *The LST formulas $\text{Const}(x, u)$ and $\text{PFml}(x, u)$ are Σ_0 .*
- (ii) *The LST formula $\text{Fml}(x, u)$ is Δ_1^{BS} . \square*

Our next task is to write down an LST formula $\text{Fr}(\varphi, x)$ such that

$$\text{Fr}(\varphi, x) \leftrightarrow \text{Fml}(\varphi) \wedge [x \text{ is the set of variables occurring free in } \varphi].$$

Now, given a formula φ , how would one go about checking whether a set x is the set of free variables of φ ? One way would be to concentrate on a sequence ψ for which $\text{Build}(\varphi, \psi)$, and proceed along the members of ψ , keeping track of the free variables at each stage. This approach leads to the following formula, which we take as our $\text{Fr}(\varphi, x)$:

$$\begin{aligned} \exists \psi \exists f [& \text{Build}(\varphi, \psi) \wedge \text{Finseq}(f) \wedge (\text{dom}(f) = \text{dom}(\psi)) \wedge (x = f(\|f\|)) \\ & \wedge (\forall i \in \text{dom}(f)) [(\exists j, k \in i) [F_\wedge(\psi_i, \psi_j, \psi_k) \wedge (f(i) = f(j) \cup f(k))] \\ & \vee (\exists j \in i) [F_\neg(\psi_i, \psi_j) \wedge (f(i) = f(j))] \\ & \vee (\exists j \in i) (\exists u \in \text{ran}(\varphi)) [\text{Vbl}(u) \wedge F_\exists(\varphi_i, u, \psi_j) \wedge (f(i) = f(j) - \{u\})] \\ & \vee [\text{PFml}(\psi_i) \wedge [[\text{Vbl}(\psi_{i_2}) \wedge \text{Vbl}((\psi_{i_3})) \wedge f(i) = \{(\psi_{i_2}), (\psi_{i_3})\}] \\ & \vee [\text{Vbl}((\psi_{i_2})) \wedge \text{Const}((\psi_{i_3})) \wedge f(i) = \{(\psi_{i_2})\}] \\ & \vee [\text{Const}((\psi_{i_2})) \wedge \text{Vbl}((\psi_{i_3})) \wedge f(i) = \{(\psi_{i_3})\}] \\ & \vee [\text{Const}((\psi_{i_2})) \wedge \text{Const}((\psi_{i_3})) \wedge f(i) = \emptyset]]]]. \end{aligned}$$

Clearly, the LST formula $\text{Fr}(\varphi, x)$ is Σ_1 .

9.8 Lemma. *The LST formula $\text{Fr}(\varphi, x)$ is Δ_1^{BS} .*

Proof. Clearly,

$$\text{BS} \vdash \text{Fr}(\varphi, x) \leftrightarrow [\text{Fml}(\varphi) \wedge \forall z [\text{Fr}(\varphi, z) \rightarrow z = x]].$$

This gives us a Π_1^{BS} characterisation of $\text{Fr}(\varphi, x)$. \square

We now formulate an LST formula $\text{Sub}(\varphi', \varphi, v, t)$, which will say that φ' and φ are formulas of \mathcal{L}_V , v is a variable, t is a constant, and φ' is the result of substituting t for every free occurrence of v in φ . To arrive at this formula, we adopt a procedure similar to the one used above for $\text{Fr}(\varphi, x)$. Pick a sequence ψ such that $\text{Build}(\varphi, \psi)$. Proceed through ψ , substituting t for every free occurrence of v at each stage. If the quantifier v is ever encountered, delete any substitutions made previously within the scope of this quantifier. In order to write this out in an intelligible fashion, we consider first the restriction of $\text{Sub}(\varphi', \varphi, v, t)$ to primitive formulas φ . Let $S(\varphi', \varphi, v, t)$ denote this restricted formula. That is, let $S(\varphi', \varphi, v, t)$ be the following LST formula:

$$\begin{aligned} & \text{PFml}(\varphi') \wedge \text{PFml}(\varphi) \wedge \text{Vbl}(v) \wedge \text{Const}(t) \\ & \wedge [[F_=(\varphi, \varphi_2, \varphi_3) \wedge [[\varphi_2 \neq v \wedge \varphi_3 \neq v \wedge (\varphi' = \varphi)] \\ & \quad \vee [\varphi_2 = v \wedge \varphi_3 \neq v \wedge F_=(\varphi', t, \varphi_3)] \\ & \quad \vee [\varphi_2 \neq v \wedge \varphi_3 = v \wedge F_=(\varphi', \varphi_2, t)] \\ & \quad \vee [\varphi_2 = v \wedge \varphi_3 = v \wedge F_=(\varphi', t, t)]]] \\ & \vee [F_\epsilon(\varphi, \varphi_2, \varphi_3) \wedge [\dots \dots \dots]], \end{aligned}$$

where the expression denoted $\dots \dots \dots$ in the above is just as in the $F_=(\varphi, \varphi_2, \varphi_3)$ part, but with F_ϵ in place of $F_=(\varphi, \varphi_2, \varphi_3)$.

Notice that $S(\varphi', \varphi, v, t)$ is Σ_0 . Let $\text{Sub}(\varphi', \varphi, v, t)$ be the following LST formula:

$$\begin{aligned} & \text{Fml}(\varphi') \wedge \text{Fml}(\varphi) \wedge \text{Vbl}(v) \wedge \text{Const}(t) \wedge \exists \psi \exists \theta [\text{Build}(\varphi, \psi) \\ & \wedge \text{Finseq}(\theta) \wedge (\text{dom}(\theta) = \text{dom}(\psi)) \wedge (\theta_{\parallel \theta \parallel} = \varphi') \\ & \wedge (\forall i \in \text{dom}(\psi)) [(\exists j, k \in i) (F_\wedge(\psi_i, \psi_j, \psi_k) \wedge F_\wedge(\theta_i, \theta_j, \theta_k)) \\ & \quad \vee (\exists j \in i) (F_\neg(\psi_i, \psi_j) \wedge F_\neg(\theta_i, \theta_j)) \\ & \quad \vee (\exists j \in i) (\exists u \in \text{ran}(\varphi)) (\text{Vbl}(u) \wedge (u \neq v) \\ & \quad \quad \wedge F_\epsilon(\psi_i, u, \psi_j) \wedge F_\exists(\theta_i, u, \theta_j)) \\ & \quad \vee (\exists j \in i) (F_\exists(\psi_i, v, \psi_j) \wedge (\theta_i = \psi_i)) \\ & \quad \vee S(\theta_i, \psi_i, v, t)]]]. \end{aligned}$$

This formula is clearly Σ_1 . Moreover:

9.9 Lemma. *The LST formula $\text{Sub}(\varphi', \varphi, v, t)$ is Δ_1^{BS} .*

Proof. Clearly,

$$\text{BS} \vdash \text{Sub}(\varphi', \varphi, v, t) \leftrightarrow \text{Fml}(\varphi) \wedge \text{Vbl}(v) \wedge \text{Const}(t) \\ \wedge \forall \psi [\text{Sub}(\psi, \varphi, v, t) \rightarrow \psi = \varphi'],$$

which gives us the lemma at once, since the expression on the right of the above equivalence is clearly Π_1^{BS} . \square

We are now able to define the notion of satisfaction (“truth”) for the languages \mathcal{L}_u . We shall write down an LST formula $\text{Sat}(u, \varphi)$ such that

$\text{Sat}(u, \varphi) \leftrightarrow \varphi$ is a sentence of \mathcal{L}_u which is true in the structure $\langle u, \in \rangle$ under the canonical interpretation (i.e. with x interpreting \hat{x} for each x in u).

The standard way to define satisfaction is as follows. Let f be a function with domain ω such that $f(0)$ is the set of all primitive formulas of \mathcal{L}_u and, in general, $f(i+1)$ is the set of all formulas of \mathcal{L}_u which are obtained from formulas in $f(i)$ by a single application of one of the three formula building schemas. Let g be a function with domain ω such that $g(i)$ is the set of all formulas in $f(i)$ which have no free variables and which are true in $\langle u, \in \rangle$. Both f and g can be defined by simple recursions. The function g then provides us with all the sentences of \mathcal{L}_u which are true in $\langle u, \in \rangle$. (The function f is required in order to handle negation in passing from $g(i)$ to $g(i+1)$.) Our formula $\text{Sat}(u, \varphi)$ will be obtained by considering the above process taken sufficiently far to check whether φ is in $g(i)$ or not, when i is chosen so that $\varphi \in f(i)$.

Let $E(\varphi, u)$ be the following LST formula:

$$(\exists x, y \in u)[(x \in y) \wedge F_e(\varphi, \hat{x}, \hat{y})] \vee (\exists x \in u) F_=(\varphi, \hat{x}, \hat{x}).$$

Clearly, $E(\varphi, u)$ says that φ is a primitive sentence of \mathcal{L}_u which is true in the structure $\langle u, \in \rangle$. Provided that we are careful when we write it out in LST, the formula $E(\varphi, u)$ is Σ_0 . For example, in rendering the clause $(\exists x \in u) F_=(\varphi, \hat{x}, \hat{x})$ in LST we must proceed thus:

$$(\exists x \in u)(\exists y \in \text{ran}(\varphi))(y = \hat{x} \wedge F_=(\varphi, y, y)).$$

It should now be clear that $E(\varphi, u)$ is Σ_0 .

The following LST formula, $S(u, \varphi)$, expresses in LST the notion that φ is a sentence of \mathcal{L}_u which is true in $\langle u, \in \rangle$. (However, as $S(u, \varphi)$ will not be Σ_1 , this is not our sought after formula $\text{Sat}(u, \varphi)$, but rather a precursor to it.)

$$(u \neq \emptyset) \wedge \text{Fml}(\varphi, u) \wedge \exists f \exists g [\text{Finseq}(f) \wedge \text{Finseq}(g) \wedge (\text{dom}(f) = \text{dom}(g)) \\ \wedge (\varphi \in g(\|g\|)) \wedge \forall \psi (\psi \in f(0) \leftrightarrow \text{PFml}(\psi, u)) \wedge \forall \psi (\psi \in g(0) \leftrightarrow E(\psi, u)) \\ \wedge (\forall j \in \text{dom}(f)) (\forall i \in j) (\forall \psi) [\psi \in f(i+1) \leftrightarrow (\psi \in f(i)) \vee (\exists \theta, \theta' \in f(i)) F_\wedge(\psi, \theta, \theta') \\ \vee (\exists \theta \in f(i)) F_\neg(\psi, \theta) \vee (\exists \theta \in f(i)) (\exists v \in \text{ran}(\psi)) (\text{Vbl}(v) \wedge F_e(\psi, v, \theta))] \wedge$$

$$\begin{aligned}
& (\forall j \in \text{dom}(g)) (\forall i \in j) (\forall \psi) [\psi \in g(i+1) \leftrightarrow (\psi \in g(i)) \\
& \quad \vee (\exists \theta, \theta' \in g(i)) F_\wedge(\psi, \theta, \theta') \vee (\exists \theta \in f(i)) (\theta \notin g(i) \wedge F_\neg(\psi, \theta)) \\
& \quad \vee (\exists \theta \in f(i)) (\exists v \in \text{ran}(\psi)) (\exists x \in u) (\exists \theta' \in g(i)) [\text{Vbl}(v) \wedge F_\exists(\psi, v, \theta) \\
& \quad \wedge \text{Sub}(\theta', \theta, v, \dot{x})]]].
\end{aligned}$$

It is easily seen that the above formula does define the satisfaction relation. But it is not a Σ_1 formula. The problem is the quantifier $(\forall \psi)$, which appears four times, and the unbounded quantifiers involved in the Δ_1^{BS} formula $\text{Sub}(\theta', \theta, v, \dot{x})$, which occurs inside the scope of a number of other quantifiers. However, it is easily seen that the truth of $S(u, \varphi)$ is not affected by binding all unbounded quantifiers involved (including the $\exists f$ and the $\exists g$) by the set

$$\begin{aligned}
w(u, \varphi) = & \left(\bigcup_{m \in \text{dom}(\varphi)}^{m+1} [\mathfrak{g} \cup \{v_i \mid i \in \omega\} \cup \{\dot{x} \mid x \in u\}] \right) \\
& \cup \left(\bigcup_{n \in \text{dom}(\varphi)}^{n+1} \left[\bigcup_{m \in \text{dom}(\varphi)}^{m+1} [\mathfrak{g} \cup \{v_i \mid i \in \omega\} \cup \{\dot{x} \mid x \in u\}] \right] \right).
\end{aligned}$$

(The first set in the above union includes all \mathcal{L}_u formulas of lengths at most that of φ , and the second set includes all finite sequences of such formulas whose domain is at most $\text{dom}(\varphi)$.) Let $S'(u, \varphi, w)$ be the formula obtained from $S(u, \varphi)$ by binding all quantifiers not already bound by w . Then for $\text{Sat}(u, \varphi)$ we take the following LST formula:

$$\begin{aligned}
& \exists w \exists x \exists y \exists a \exists b \exists t [(a = \{\dot{x} \mid x \in u\}) \wedge ("t = \omega") \wedge (b = \{v_i \mid i \in t\}) \\
& \quad \wedge \text{Seq}(x, \mathfrak{g} \cup a \cup b, \text{dom}(\varphi) + 1) \wedge \text{Seq}(y, x, \text{dom}(\varphi) + 1) \\
& \quad \wedge (w = x \cup y) \wedge S'(u, \varphi, w)],
\end{aligned}$$

where the formula " $t = \omega$ " is written out thus:

$$\text{On}(t) \wedge \text{lim}(t) \wedge (\forall i \in t) [(\exists j \in i)(i = j + 1) \vee (\forall j \in i)(j \neq j)].$$

By our previous remarks, $\text{Sat}(u, \varphi)$ is equivalent to $S(u, \varphi)$, so indeed

$$\text{Sat}(u, \varphi) \leftrightarrow \varphi \text{ is a sentence of } \mathcal{L}_u \text{ which is true in } \langle u, \in \rangle.$$

Moreover, $\text{Sat}(u, \varphi)$ is clearly Σ_1 , and in fact:

9.10 Lemma. *The LST formula $\text{Sat}(u, \varphi)$ is Δ_1^{BS} .*

Proof. Clearly

$$\text{BS} \vdash \neg \text{Sat}(u, \varphi) \leftrightarrow \neg [\text{Fml}(\varphi, u) \wedge \text{Fr}(\varphi, \emptyset)] \vee \exists \theta [F_\neg(\theta, \varphi) \wedge \text{Sat}(u, \theta)].$$

Hence $\neg \text{Sat}(u, \varphi)$ is Σ_1^{BS} , whence $\text{Sat}(u, \varphi)$ is Π_1^{BS} . \square

We often write $\vDash_u \varphi$ instead of $\text{Sat}(u, \varphi)$.

As we have remarked earlier, the collection of sets which constitute the "formulas" of \mathcal{L} provides us with an analogue of the formulas of the genuine language

LST. Given any formula Φ of LST we can construct a set φ which, according to the “syntax” of \mathcal{L} developed above, has the same logical structure as Φ . In this context, the following result indicates how the formal notion of satisfaction just defined corresponds to the genuine notion of truth.

9.11 Lemma. *Let $\Phi(v_0, \dots, v_n)$ be any formula of LST, and let $\varphi(v_0, \dots, v_n)$ be its counterpart in \mathcal{L} (in the sense described above). Then*

$$\text{ZF} \vdash \forall u (\forall x_0 \in u) \dots (\forall x_n \in u) [\Phi^u(x_0, \dots, x_n) \leftrightarrow \text{Sat}(u, \varphi(\dot{x}_0, \dots, \dot{x}_n))].$$

Proof. By induction on the construction of Φ and φ . (The easy details are left as an exercise for the reader.) \square

Notice that the above result is a theorem schema for ZF, which takes us from a given LST formula Φ and the genuine notion of truth to a “formula” φ of \mathcal{L} and the mathematically defined notion of satisfaction.

By analogy with LST, we define a “Lévy hierarchy” for the formulas of \mathcal{L}_V . For reasons of technical convenience we only allow for single quantifiers rather than blocks of like quantifiers as we did for LST.

A formula φ of \mathcal{L}_V is said to be Σ_0 (or Π_0) if, whenever a quantifier $\exists v_n$ occurs in φ it does so in the context

$$\exists v_n (v_n \in x \wedge \dots)$$

for some $x \in \text{Vbl} \cup \text{Const}$. The following LST formula, $\text{Fml}^{\Sigma_0}(\varphi)$, clearly defines this notion:

$$\begin{aligned} \text{Fml}(\varphi) \wedge (\forall i \in \text{dom}(\varphi)) [& (\varphi_i = 0 \wedge \varphi_{i+1} = 8 \wedge \text{Vbl}(\varphi_{i+2})) \\ & \rightarrow (\varphi_{i+3} = 0 \wedge \varphi_{i+4} = 6 \wedge \varphi_{i+5} = 0 \wedge \varphi_{i+6} = 4 \\ & \wedge \varphi_{i+7} = \varphi_{i+2} \wedge (\text{Const}(\varphi_{i+8}) \vee \text{Vbl}(\varphi_{i+8})) \wedge \varphi_{i+9} = 1]. \end{aligned}$$

Notice that except for the part $\text{Fml}(\varphi)$, this formula is Σ_0 . Likewise for the LST formula $\text{Fml}^{\Sigma_0}(\varphi, u)$, which says that φ is a Σ_0 formula of \mathcal{L}_u . The following lemma is immediate:

9.12 Lemma. *The LST formulas $\text{Fml}^{\Sigma_0}(\varphi)$ and $\text{Fml}^{\Sigma_0}(\varphi, u)$ are Δ_1^{BS} . \square*

A formula φ of \mathcal{L}_V is said to be Σ_1 if it is of the form $\exists v_n \psi$, where ψ is Σ_0 , and is said to be Π_1 if it is of the form $\neg \exists v_n \psi$ where ψ is Σ_0 . In general, an \mathcal{L}_V formula is said to be Σ_{n+1} if it is of the form $\exists v_n \psi$ where ψ is Π_n , and is said to be Π_{n+1} if it is of the form $\neg \psi$ where ψ is Σ_{n+1} .

9.13 Lemma. *Fix $n \geq 1$. Then there are Δ_1^{BS} formulas $\text{Fml}^{\Sigma_n}(\varphi)$, $\text{Fml}^{\Pi_n}(\varphi)$, $\text{Fml}^{\Sigma_n}(\varphi, u)$, $\text{Fml}^{\Pi_n}(\varphi, u)$ of LST such that:*

$$\begin{aligned} \text{Fml}^{\Sigma_n}(\varphi) &\leftrightarrow \varphi \text{ is a } \Sigma_n \text{ formula of } \mathcal{L}_V; \\ \text{Fml}^{\Pi_n}(\varphi) &\leftrightarrow \varphi \text{ is a } \Pi_n \text{ formula of } \mathcal{L}_V; \\ \text{Fml}^{\Sigma_n}(\varphi, u) &\leftrightarrow \varphi \text{ is a } \Sigma_n \text{ formula of } \mathcal{L}_u; \\ \text{Fml}^{\Pi_n}(\varphi, u) &\leftrightarrow \varphi \text{ is a } \Pi_n \text{ formula of } \mathcal{L}_u. \end{aligned}$$

Proof. These are all more or less the same as $\text{Fml}^{\Sigma_0}(\varphi)$, considered earlier. For example, $\text{Fml}^{\Sigma_1}(\varphi)$ is:

$$\begin{aligned} & \text{Fml}(\varphi) \wedge [\varphi_0 = 0 \wedge \varphi_1 = 8 \wedge \text{Vbl}(\varphi_2) \\ & \quad \wedge (\forall i \in \text{dom}(\varphi))[(i > 2 \wedge \varphi_i = 0 \wedge \varphi_{i+1} = 8 \wedge \text{Vbl}(\varphi_{i+2})) \\ & \quad \rightarrow (\varphi_{i+3} = 0 \wedge \varphi_{i+4} = 6 \wedge \varphi_{i+5} = 0 \wedge \varphi_{i+6} = 4 \wedge \varphi_{i+7} = \varphi_{i+2} \\ & \quad \wedge (\text{Const}(\varphi_{i+8}) \vee \text{Vbl}(\varphi_{i+8})) \wedge \varphi_{i+9} = 1]]]. \end{aligned}$$

(As n increases, the length and complexity of $\text{Fml}^{\Sigma_n}(\varphi)$, etc. also increases, of course, but the overall pattern is much the same.) \square

Occasionally we shall wish to consider extensions of the languages \mathcal{L}_u in which there are a finite number of additional predicates. Specifically, let k be some natural number and let $A_1 \subseteq u^{n(1)}, \dots, A_k \subseteq u^{n(k)}$. The language $\mathcal{L}_u(\mathring{A}_1, \dots, \mathring{A}_k)$ has the same structure as \mathcal{L}_u except that there are the k extra predicate letters $\mathring{A}_1, \dots, \mathring{A}_k$, where \mathring{A}_i is $n(i)$ -ary for each i . More precisely, for each $i = 1, \dots, k$, amongst the primitive formulas of $\mathcal{L}_u(\mathring{A}_1, \dots, \mathring{A}_k)$ we allow the sequences

$$\langle 0, 8 + i, x_1, \dots, x_{n(i)}, 1 \rangle,$$

where $x_1, \dots, x_{n(i)} \in \text{Vbl} \cup \text{Const}_u$. We usually write $\mathring{A}_i(x_1, \dots, x_{n(i)})$ in place of the sequence $\langle 0, 8 + i, x_1, \dots, x_{n(i)}, 1 \rangle$. With this modification to the primitive formulas, the development of the rest of the language $\mathcal{L}_u(\mathring{A}_1, \dots, \mathring{A}_k)$ proceeds exactly as for \mathcal{L}_u . Consequently, all of the results obtained in this section for the languages \mathcal{L}_u hold in this more general situation. (Note that the interpretation of $\mathcal{L}_u(\mathring{A}_1, \dots, \mathring{A}_k)$ in the structure $\langle u, \in, A_1, \dots, A_k \rangle$ is the obvious, canonical one.)

We shall require the following formal analogue of the metamathematical notion of absoluteness (Section 8).

Let $\varphi(\vec{x})$ be any $\mathcal{L}(\mathring{A}_1, \dots, \mathring{A}_k)$ formula. Let \mathbf{M}, \mathbf{N} be structures appropriate for this language, \mathbf{M} a substructure of \mathbf{N} . We say that φ is *U-absolute* for \mathbf{M}, \mathbf{N} iff

$$(\forall \vec{x} \in M)(\vDash_{\mathbf{M}} \varphi(\vec{x}) \text{ implies } \vDash_{\mathbf{N}} \varphi(\vec{x})).$$

We say that φ is *D-absolute* for \mathbf{M}, \mathbf{N} iff

$$(\forall \vec{x} \in M)(\vDash_{\mathbf{N}} \varphi(\vec{x}) \text{ implies } \vDash_{\mathbf{M}} \varphi(\vec{x})).$$

We say that φ is *absolute* for \mathbf{M}, \mathbf{N} iff it is both *U-absolute* and *D-absolute* for \mathbf{M}, \mathbf{N} . Analogous to 8.3 we have:

9.4 Lemma. *Let \mathbf{M}, \mathbf{N} be $\mathcal{L}(\mathring{A}_1, \dots, \mathring{A}_k)$ structures, \mathbf{M} a substructure of \mathbf{N} . Suppose further that both M and N are transitive sets. Let $\varphi(\vec{x})$ be a formula of $\mathcal{L}(\mathring{A}_1, \dots, \mathring{A}_k)$.*

- (i) *If φ is Σ_0 , then φ is absolute for \mathbf{M}, \mathbf{N} .*
- (ii) *If φ is Σ_1 , then φ is U-absolute for \mathbf{M}, \mathbf{N} .*
- (iii) *If φ is Π_1 , then φ is D-absolute for \mathbf{M}, \mathbf{N} .*

Proof. Similar to the proof of 8.3. (The details are left as an exercise for the reader.) \square

The following lemma concerns the relationship between the two languages LST and \mathcal{L} , and is related to Lemma 9.11.

9.15 Lemma. *Let $\Phi(\vec{x})$ be a Σ_0 formula of LST, and let $\varphi(\vec{x})$ be its counterpart in \mathcal{L} . Then*

$$\text{ZF} \vdash \text{“For any transitive set } M, (\forall \vec{x} \in M)[\Phi(\vec{x}) \leftrightarrow \vDash_M \varphi(\vec{x})]\text{”}.$$

Proof. By an easy induction on the length of Φ . (The details are left as an exercise to the reader.) \square

We shall make considerable use of 9.15 and generalisations thereof in Chapter II.

10. Definability

Consider a structure of the form

$$\mathbf{M} = \langle M, \in, A_1, \dots, A_k \rangle,$$

where M is a non-empty set and $A_i \subseteq M^{n(i)}$ for $i = 1, \dots, k$. (In such cases we often omit specific reference to \in , as is always the case with $=$, of course.) By the **M-language** we mean the language $\mathcal{L}_M(\dot{A}_1, \dots, \dot{A}_k)$ introduced at the end of the previous section. As we indicated there, all of the various definitions and results of section 9 hold for **M-languages**. For instance, there is a Δ_1^{BS} formula $\text{Sat}(\mathbf{M}, \varphi)$ of LST such that $\text{Sat}(\mathbf{M}, \varphi)$ iff φ is a sentence of the **M-language** which is true in \mathbf{M} (under the standard interpretation). Note that we usually write $\vDash_{\mathbf{M}} \varphi$ instead of $\text{Sat}(\mathbf{M}, \varphi)$.

Let $N \subseteq M$. A set $R \subseteq M^m$ is said to be $\Sigma_n^{\mathbf{M}}(N)$ iff there is a Σ_n formula $\varphi(v_0, \dots, v_{m-1})$ of the **M-language**, whose constants are all members of the set $\{\dot{a} \mid a \in N\}$, such that

$$(\forall x_0, \dots, x_{m-1} \in M)[(x_0, \dots, x_{m-1}) \in R \leftrightarrow \vDash_{\mathbf{M}} \varphi(\dot{x}_0, \dots, \dot{x}_{m-1})].$$

Similarly for $\Pi_n^{\mathbf{M}}(N)$. A set $R \subseteq M^m$ is $\Delta_n^{\mathbf{M}}(N)$ iff it is both $\Sigma_n^{\mathbf{M}}(N)$ and $\Pi_n^{\mathbf{M}}(N)$.

We write $\Sigma_n^{\mathbf{M}}$ instead of $\Sigma_n^{\mathbf{M}}(\emptyset)$ and $\Sigma_n(\mathbf{M})$ instead of $\Sigma_n^{\mathbf{M}}(M)$. Similarly for Π and Δ .

A set $R \subseteq M^m$ is said to be **M-definable** iff it is $\Sigma_n(\mathbf{M})$ for some n .

Notice that the above notions are all formally defined within set theory, and are not metamathematical notions. For example, there is an LST formula $\Phi(R, M)$ such that

$$\Phi(R, M) \leftrightarrow M \text{ is a non-empty set} \wedge R \subseteq M \wedge R \text{ is } M\text{-definable}.$$

(As an exercise, the reader may like to investigate the logical complexity of such a formula.)

It φ is a formula of the \mathbf{M} -language, the interpretations (in \mathbf{M}) of the constants \dot{x} which occur in φ are called the *parameters* of φ .

Let A be some class of m -tuples, \mathbf{M} a given structure. We say that the class A is $\Sigma_n^{\mathbf{M}}(N)$ iff $A \cap M^m$ is $\Sigma_n^{\mathbf{M}}(N)$, etc.

A related notion is the following. Let \mathcal{F} be a class of structures of the form $\mathbf{M} = \langle M, \in, A_1, \dots, A_k \rangle$, where k is fixed and each A_i is $n(i)$ -ary, for fixed $n(i)$, $i = 1, \dots, k$. Let A be a class of m -tuples. We say that A is *uniformly* $\Sigma_n^{\mathbf{M}}$ for $\mathbf{M} \in \mathcal{F}$ iff there is a *single* Σ_n formula $\varphi(v_0, \dots, v_{m-1})$ of $\mathcal{L}(\dot{A}_1, \dots, \dot{A}_k)$ such that for each $\mathbf{M} \in \mathcal{F}$,

$$A \cap M^m = \{(x_0, \dots, x_{m-1}) \mid \vDash_{\mathbf{M}} \varphi(\dot{x}_0, \dots, \dot{x}_{m-1})\}.$$

Similar for *uniformly* $\Pi_n^{\mathbf{M}}$ and *uniformly* $\Delta_n^{\mathbf{M}}$. We shall presently give some examples of these important (to us) concepts. In order to do so, however, we require some preliminary ideas.

A set M is said to be *amenable* iff it is transitive and satisfies the following conditions:

- (i) $(\forall x, y \in M)(\{x, y\} \in M)$;
- (ii) $(\forall x \in M)(\bigcup x \in M)$;
- (iii) $\omega \in M$;
- (iv) $(\forall x, y \in M)(x \times y \in M)$;
- (v) if $R \subseteq M$ is $\Sigma_0(M)$, then $(\forall x \in M)(R \cap x \in M)$.

(Intuitively speaking, an amenable set is thus a transitive “model” of the theory BS of section 9. The idea behind this definition is that it will enable us to prove, within set theory, semantic analogues of the logical complexity results of section 9.)

Notice that if M is amenable, then $x \in M$ whenever $x \subseteq M$ is finite.

10.1 Lemma. *The predicate “ x is finite” is uniformly Σ_1^M for amenable M .*

Proof. Let $\Phi(x, n, f)$ be the Σ_0 LST formula

$$(n \text{ is a natural number}) \wedge (f: n \leftrightarrow x).$$

Clearly, for any set x ,

$$x \text{ is finite} \leftrightarrow \exists n \exists f \Phi(x, n, f).$$

Let φ be the analogue to Φ in \mathcal{L} . We prove that for any amenable set M ,

$$(\forall x \in M)[\exists n \exists f \Phi(x, n, f) \leftrightarrow \vDash_M \exists n \exists f \varphi(\dot{x}, n, f)],$$

which proves the lemma, of course.

Let M be amenable, $x \in M$. Suppose first that

$$\vDash_M \exists n \exists f \varphi(\dot{x}, n, f).$$

Then by 9.11,

$$[\exists n \exists f \Phi(x, n, f)]^M.$$

But Σ_1 formulas of LST are U -absolute for transitive classes. Hence,

$$\exists n \exists f \Phi(x, n, f),$$

as required. Now suppose that this last formula is true. Pick n, f such that

$$\Phi(x, n, f).$$

Since M is transitive and $\omega \in M$ we have $\omega \subseteq M$, so certainly $n \in M$. Hence $x \times n \in M$. But $f \subseteq x \times n$ and f is finite, so $f \in M$. Then by the D -absoluteness of all Σ_0 formulas of LST, we have

$$[\Phi(x, n, f)]^M.$$

Thus

$$[\exists n \exists f \Phi(x, n, f)]^M,$$

and by 9.11 we conclude that

$$\vDash_M \exists n \exists f \varphi(\dot{x}, n, f),$$

and we are done. \square

If $R(x_0, \dots, x_m)$ and $S(x_0, \dots, x_m)$ are relations on M , then, extending our convention that $R(x_0, \dots, x_m)$ means $(x_0, \dots, x_m) \in R$, etc., we write:

$$\begin{array}{ll} (R \wedge S)(x_0, \dots, x_m) & \text{iff } (x_0, \dots, x_m) \in R \cap S, \\ (R \vee S)(x_0, \dots, x_m) & \text{iff } (x_0, \dots, x_m) \in R \cup S, \\ (\neg R)(x_0, \dots, x_m) & \text{iff } (x_0, \dots, x_m) \in M^{m+1} - R, \\ (\exists x_0 R)(x_1, \dots, x_m) & \text{iff } (\exists x_0 \in M)((x_0, x_1, \dots, x_m) \in R), \\ ((\exists x_0 \in z) R)(z, x_1, \dots, x_m) & \text{iff } (\exists x_0 \in z)((x_0, x_1, \dots, x_m) \in R), \\ \text{etc.} & \end{array}$$

By means of, in particular, quantifier contraction along the lines of 8.9, we can easily prove:

10.2 Lemma. *Let M be an amenable set, and let $\mathbf{M} = \langle M, A_1, \dots, A_k \rangle$. Let R, S be m -ary relations on M .*

- (i) *If R, S are $\Sigma_0^{\mathbf{M}}(N)$, so too are $R \wedge S, R \vee S, \neg R$.*
- (ii) *If R is $\Sigma_n^{\mathbf{M}}(N)$, $\neg R$ is $\Pi_n^{\mathbf{M}}(N)$.*
- (iii) *If R is $\Pi_n^{\mathbf{M}}(N)$, $\neg R$ is $\Sigma_n^{\mathbf{M}}(N)$.*
- (iv) *R is $\Delta_n^{\mathbf{M}}(N)$ iff both R and $\neg R$ are $\Sigma_n^{\mathbf{M}}(N)$.*
- (v) *If R, S are $\Sigma_{n+1}^{\mathbf{M}}(N)$, so are $R \wedge S, R \vee S, \exists x R, (\exists x \in z) R$.*
- (vi) *If R, S are $\Pi_{n+1}^{\mathbf{M}}(N)$, so are $R \wedge S, R \vee S, \forall x R, (\forall x \in z) R$.*
- (vii) *If R, S are $\Delta_{n+1}^{\mathbf{M}}(N)$, so are $R \wedge S, R \vee S, \neg R$. \square*

The following simple lemma employs the same trick used in the proofs of both 9.8 and 9.9.

10.3 Lemma. *Let M be an amenable set, and let $\mathbf{M} = \langle M, A_1, \dots, A_k \rangle$. Let $f \subseteq M \times M$ be a function. (We say f is a function over M in this case.) If f is $\Sigma_n^{\mathbf{M}}(N)$ and $\text{dom}(f)$ is $\Pi_n^{\mathbf{M}}(N)$, then both f and $\text{dom}(f)$ are $\Delta_n^{\mathbf{M}}(N)$.*

Proof. Since

$$x \in \text{dom}(f) \leftrightarrow \exists y [y = f(x)],$$

we see at once that $\text{dom}(f)$ is $\Sigma_n^{\mathbf{M}}(N)$. To see that f is $\Pi_n^{\mathbf{M}}(N)$, note the equivalence

$$y = f(x) \leftrightarrow [x \in \text{dom}(f)] \wedge \forall z [z = f(x) \rightarrow y = z]. \quad \square$$

10.4 Corollary. *Let \mathbf{M} be as above. If $f: M \rightarrow M$ is $\Sigma_n^{\mathbf{M}}(N)$, then f is in fact $\Delta_n^{\mathbf{M}}(N)$. \square*

10.5 Lemma. *Let M be amenable, and let $\mathbf{M} = \langle M, A_1, \dots, A_k \rangle$. Let $n \geq 1$ and let f be a $\Sigma_n^{\mathbf{M}}(N)$ m -ary function over M (i.e. $f \subseteq M^{m+1}$). Let g be a $\Sigma_n^{\mathbf{M}}(N)$ unary function over M and let R be a $\Sigma_n^{\mathbf{M}}(N)$ unary relation on M . Then h, S are $\Sigma_n^{\mathbf{M}}(N)$, where:*

(i) h is the m -ary function defined by

$$h(\vec{x}) = g \circ f(\vec{x});$$

(ii) S is the m -ary relation defined by

$$S(\vec{x}) \leftrightarrow R(f(\vec{x})).$$

Proof. By 10.2 and the observations

$$\begin{aligned} y = h(\vec{x}) &\leftrightarrow \exists z [y = g(z) \wedge z = f(\vec{x})], \\ S(\vec{x}) &\leftrightarrow \exists z [R(z) \wedge z = f(\vec{x})]. \quad \square \end{aligned}$$

10.6 Lemma. *Let M be amenable, and let $\mathbf{M} = \langle M, A_1, \dots, A_k \rangle$. If $R(x)$ is a $\Sigma_n^{\mathbf{M}}(N)$ unary relation on M , so too is $Q(x)$, where*

$$Q(x) \leftrightarrow [x \text{ is an ordered pair} \wedge R((x)_0)].$$

Similarly for $(x)_1$, etc. (We usually write $R((x)_0)$ in place of $Q(x)$ as defined above, etc.)

Proof. $Q(x) \leftrightarrow x \text{ is an ordered pair} \wedge (\exists u \in x)(\exists y \in u)(y = (x)_0 \wedge R(y))$. \square

10.7 Lemma (Contraction of Parameters). *Let M be as above. Let $n \geq 1$, and let R be a $\Sigma_n(\mathbf{M})$ relation on M . Then there is a single element $p \in M$ such that R is $\Sigma_n^{\mathbf{M}}(\{p\})$.*

Proof. Let R be $\Sigma_n^{\mathbf{M}}(\{p_1, \dots, p_m\})$. Set

$$p = (p_1, \dots, p_m).$$

Using the method of 10.6 it is easily seen that R is $\Sigma_n^{\mathbf{M}}(\{p\})$. \square

We note also the following consequence of 9.11:

10.8 Lemma. *Fix $n \geq 1$. Let M be amenable, and let $\mathbf{M} = \langle M, A_1, \dots, A_k \rangle$. If $R(\vec{z})$ is a $\Sigma_n^{\mathbf{M}}(N)$ relation on M , there is a $\Sigma_0^{\mathbf{M}}(N)$ relation $S(\vec{x}, \vec{z})$ on M such that*

$$R(\vec{z}) \leftrightarrow (\exists x_1 \in M)(\forall x_2 \in M)(\exists x_3 \in M) \dots (-x_n \in M) S(\vec{x}, \vec{z}). \quad \square$$

For later use we make the following definitions. Let $\mathbf{M} = \langle M, A_1, \dots, A_k \rangle$, $\mathbf{N} = \langle N, B_1, \dots, B_k \rangle$. We say that \mathbf{N} is an *elementary substructure* of \mathbf{M} , and write $\mathbf{N} \prec \mathbf{M}$, iff $N \subseteq M$, B_i is the restriction of A_i to N for $i = 1, \dots, k$, and for all sentences φ of $\mathcal{L}_N(\hat{A}_1, \dots, \hat{A}_k)$.

$$\vDash_{\mathbf{N}} \varphi \quad \text{iff} \quad \vDash_{\mathbf{M}} \varphi.$$

(Notice that the sentence φ may contain constants denoting elements of N .) For $n \geq 0$, we say that \mathbf{N} is a Σ_n *elementary substructure* of \mathbf{M} , and write $\mathbf{N} \prec_n \mathbf{M}$, iff the above holds when φ is restricted to be a Σ_n sentence. We shall write $X \prec \mathbf{M}$ to mean that X is the domain of a (necessarily unique for X) elementary substructure of \mathbf{M} , and analogously $X \prec_n \mathbf{M}$. We write $\pi: \mathbf{N} \prec \mathbf{M}$ (respectively $\pi: \mathbf{N} \prec_n \mathbf{M}$) iff π is an isomorphism from \mathbf{N} to an elementary (respectively Σ_n elementary) substructure of \mathbf{M} .

11. Kripke-Platek Set Theory. Admissible Sets

We have already worked with one subtheory of ZF, namely the Basic Set Theory, BS. In this section we consider another, much stronger subtheory: *Kripke-Platek Set Theory*, KP. This is a particularly important subtheory of ZF for various reasons. One reason, of relevance to us, is that KP is the weakest subtheory of ZF which suffices for the construction of the constructible hierarchy of sets, introduced in Chapter II.

The theory KP is the LST theory whose axioms are the axioms of BS, together with the Σ_0 *Collection Schema*:

$$\forall \vec{a} [\forall x \exists y \Phi(y, x, \vec{a}) \rightarrow \forall u \exists v (\forall x \in u) (\exists y \in v) \Phi(y, x, \vec{a})],$$

where Φ is a Σ_0 formula of LST.

By an *admissible set* we mean an amenable set M (see section 10) such that for any $\Sigma_0(M)$ relation $R \subseteq M \times M$, if

$$(\forall x \in M)(\exists y \in M) R(y, x)$$

then for any $u \in M$ there is a $v \in M$ such that

$$(\forall x \in u)(\exists y \in v) R(y, x).$$

Clearly, the notion of an admissible set is related to the theory KP in the same way that the notion of an amenable set is related to the theory BS (i.e. admissible sets are transitive “models” of the theory KP.)

For κ an uncountable cardinal, we define

$$H_\kappa = \{x \mid |\text{TC}(x)| < \kappa\}.$$

Using the following lemma, we shall be able to show that H_κ is an admissible set for any uncountable cardinal κ .

11.1 Lemma. *Let $\varphi(\vec{x})$ be a Σ_1 formula of \mathcal{L} . Let κ, λ be uncountable cardinals, $\lambda < \kappa$. If $\vec{x} \in H_\lambda$ are such that $\vDash_{H_\kappa} \varphi(\vec{x})$, then $\vDash_{H_\lambda} \varphi(\vec{x})$.*

Proof. Let

$$W = \text{TC}(\{\vec{x}\}).$$

Clearly, $W \in H_\lambda$. Pick $M \prec H_\kappa$ with $W \subseteq M$ and $|M| = |W| < \lambda$. (That this can always be done follows from the Löwenheim-Skolem-Tarski Theorem. We assume the reader is familiar with this theorem.) Let

$$\pi: M \cong N$$

be the collapsing isomorphism (see 7.1), where N is transitive. Then $|N| = |M| < \lambda$, so $N \in H_\lambda$ and $N \subseteq H_\lambda$. Now, $\pi^{-1}: N \prec H_\kappa$ and (see 7.1) $\pi \upharpoonright W = \text{id} \upharpoonright W$, so $\vDash_N \varphi(\vec{x})$. But φ is Σ_1 , so by 9.14, φ is U -absolute for N, H_λ . Thus $\vDash_{H_\lambda} \varphi(\vec{x})$, as required. \square

11.2 Lemma. *If κ is an uncountable cardinal, then H_κ is admissible.*

Proof. It is easily seen that H_κ is amenable for any uncountable cardinal κ . (Exercise: Check this.) Moreover, it is also easy to see that in the case where κ is regular, H_κ is in fact admissible. We are therefore left with proving admissibility in the case where κ is singular.

So assume that κ is singular, and let $R \subseteq H_\kappa \times H_\kappa$ be $\Sigma_0(H_\kappa)$. We must show that if

$$(\forall x \in H_\kappa)(\exists y \in H_\kappa) R(y, x)$$

and if $u \in H_\kappa$, then there is a $v \in H_\kappa$ such that

$$(\forall x \in u)(\exists y \in v) R(y, x).$$

Let $\varphi(y, x, \vec{a})$ be a Σ_0 formula of \mathcal{L} and $\vec{a} \in H_\kappa$ be such that

$$R(y, x) \leftrightarrow \vDash_{H_\kappa} \varphi(\check{y}, \check{x}, \vec{\check{a}}).$$

Let $u \in H_\kappa$ be given. We seek a $v \in H_\kappa$ such that

$$\vDash_{H_\kappa} (\forall x \in \check{u})(\exists y \in \check{v}) \varphi(y, x, \vec{\check{a}}).$$

Let

$$W = \text{TC}(\{u, \vec{a}\}).$$

Then $W \in H_\kappa$, so as κ is singular there is a regular cardinal $\lambda < \kappa$ such that $W \in H_\lambda$. Now, by the assumptions on R ,

$$\vDash_{H_\kappa} \forall x \exists y \varphi(y, x, \vec{a}).$$

So for all $x \in H_\kappa$,

$$\vDash_{H_\kappa} \exists y \varphi(y, \overset{\circ}{x}, \vec{a}).$$

So by 11.1, for all $x \in H_\lambda$,

$$\vDash_{H_\lambda} \exists y \varphi(y, \overset{\circ}{x}, \vec{a}).$$

Thus

$$\vDash_{H_\lambda} \forall x \exists y \varphi(y, x, \vec{a}).$$

But λ is regular, so as we observed above, H_λ is admissible. Thus as $u \in H_\lambda$, there is a $v \in H_\lambda$ such that

$$\vDash_{H_\lambda} (\forall x \in \overset{\circ}{u})(\exists y \in \overset{\circ}{v}) \varphi(y, x, \vec{a}).$$

But the sentence involved here is Σ_0 , and hence (by 9.14) absolute for H_λ, H_κ . Thus

$$\vDash_{H_\kappa} (\forall x \in \overset{\circ}{u})(\exists y \in \overset{\circ}{v}) \varphi(y, x, \vec{a}),$$

and we are done. \square

We shall obtain a few elementary results about the theory KP. Our first two show that KP entails stronger versions of the Collection and Comprehension Axioms than were allowed for in the axioms.

11.3 Lemma (Σ_1 -Collection Principle). *Let $\Phi(y, x, \vec{a})$ be a Σ_1 formula of LST. Then*

$$\text{KP} \vdash \forall \vec{a} [\forall x \exists y \Phi(y, x, \vec{a}) \rightarrow \forall u \exists v (\forall x \in u)(\exists y \in v) \Phi(y, x, \vec{a})].$$

Proof. Let $\Psi(z, y, x, \vec{a})$ be a Σ_0 formula of LST such that

$$\text{KP} \vdash \Phi(y, x, \vec{a}) \leftrightarrow \exists z \Psi(z, y, x, \vec{a}).$$

(By 8.9, such a formula can always be found.) Argue in KP from now on.

Let \vec{a} be given, and assume

$$\forall x \exists y \Phi(y, x, \vec{a}).$$

Then

$$\forall x \exists y \exists z \Psi(z, y, x, \vec{a}).$$

Hence

$$\forall x \exists w \Psi((w)_0, (w)_1, x, \vec{a}).$$

Given u now we must find a v such that

$$(\forall x \in u)(\exists y \in v) \Phi(y, x, \vec{a}).$$

But by 8.8, the formula $\Psi((w)_0, (w)_1, x, \vec{a})$ is Σ_0 , so by Σ_0 -Collection there is a t such that

$$(\forall x \in u)(\exists w \in t) \Psi((w)_0, (w)_1, x, \vec{a}).$$

Let $v = \bigcup t$. Then

$$(\forall x \in u)(\exists y \in v)(\exists z) \Psi(z, y, x, \vec{a}).$$

Hence

$$(\forall x \in u)(\exists y \in v) \Phi(y, x, \vec{a}),$$

as required. \square

11.1 Lemma (Δ_1 -Comprehension Principle). *Let $\Phi(z, \vec{a})$ be a Δ_1^{KP} formula of LST. Then*

$$\text{KP} \vdash \forall \vec{a} \forall x \exists y \forall z [z \in y \leftrightarrow z \in x \wedge \Phi(z, \vec{a})].$$

Proof. By 8.9 we can find Σ_0 formulas Θ, Ψ of LST such that

$$\text{KP} \vdash \Phi(z, \vec{a}) \leftrightarrow \forall v \Theta(v, z, \vec{a}),$$

$$\text{KP} \vdash \Phi(z, \vec{a}) \leftrightarrow \exists v \Psi(v, z, \vec{a}).$$

We argue in KP from now on.

Let \vec{a}, x be given. We seek a y such that

$$\forall z [z \in y \leftrightarrow z \in x \wedge \Phi(z, \vec{a})].$$

Now,

$$\forall z [\Phi(z, \vec{a}) \vee \neg \Phi(z, \vec{a})].$$

Hence

$$\forall z \exists v [\Psi(v, z, \vec{a}) \vee \neg \Theta(v, z, \vec{a})].$$

By Σ_0 -Collection there is thus a set u such that

$$(*) \quad (\forall z \in x)(\exists v \in u) [\Psi(v, z, \vec{a}) \vee \neg \Theta(v, z, \vec{a})].$$

By Σ_0 -Comprehension, let

$$y = \{z \in x \mid (\exists v \in u) \Psi(v, z, \vec{a})\}.$$

We finish by showing that

$$y = \{z \in x \mid \Phi(z, \vec{a})\}.$$

Certainly, for any $z \in x$,

$$\begin{aligned} (\exists v \in u) \Psi(v, z, \vec{a}) &\rightarrow \exists v \Psi(v, z, \vec{a}) \\ &\leftrightarrow \Phi(z, \vec{a}), \end{aligned}$$

so what we must prove is that for any $z \in x$,

$$\Phi(z, \vec{a}) \rightarrow (\exists v \in u) \Psi(v, z, \vec{a}).$$

By (*), we know that there is a $v \in u$ such that

$$\Psi(v, z, \vec{a}) \vee \neg \Theta(v, z, \vec{a}).$$

But

$$\Phi(z, \vec{a}) \leftrightarrow \forall v \Theta(v, z, \vec{a}).$$

Hence for the $v \in u$ chosen above, we must have

$$\Psi(v, z, \vec{a}).$$

We are done. \square

The following lemma is a useful alternative to the Σ_1 -Collection Principle (11.3).

11.5 Lemma (Localised Σ_1 -Collection Schema). *If Φ is a Σ_1 formula of LST, then*

$$\mathbf{KP} \vdash \forall \vec{a} [(\forall x \in u) \exists y \Phi(y, x, \vec{a}) \rightarrow \exists v (\forall x \in u) (\exists y \in v) \Phi(y, x, \vec{a})].$$

Proof. Argue in \mathbf{KP} . Assume

$$(\forall x \in u) \exists y \Phi(y, x, \vec{a}).$$

Then

$$\forall x \exists y (x \notin u \vee \Phi(y, x, \vec{a})).$$

So by Σ_1 -Collection there is a v such that

$$(\forall x \in u) (\exists y \in v) (x \notin u \vee \Phi(y, x, \vec{a})).$$

But this is logically equivalent to

$$(\forall x \in u) (\exists y \in v) \Phi(y, x, \vec{a}),$$

and we are done. \square

The next lemma extends 8.6 (iv), (v) for the theory \mathbf{KP} , and is a special case of 8.10.

- 11.6 Lemma.** (i) If $\Phi(y, \vec{x})$ is a Σ_1 formula of LST, then $(\forall z \in y) \Phi(z, \vec{x})$ is Σ_1^{KP} .
(ii) If $\Phi(y, \vec{x})$ is a Π_1 formula of LST, then $(\exists z \in y) \Phi(z, \vec{x})$ is Π_1^{KP} .

Proof. We prove (i); (ii) then follows by taking negations. Let $\Psi(w, y, \vec{x})$ be a Σ_0 formula such that

$$\text{KP} \vdash \Phi(y, \vec{x}) \leftrightarrow \exists w \Psi(w, y, \vec{x}).$$

By 8.9, such a Ψ can be found, of course. We argue in KP from now on.

We have:

$$\begin{aligned} (\forall z \in y) \Phi(z, \vec{x}) &\leftrightarrow (\forall z \in y) (\exists w) \Psi(w, z, \vec{x}) \\ &\rightarrow (\exists v) (\forall z \in y) (\exists w \in v) \Psi(w, z, \vec{x}) \quad (\text{by 11.5}) \\ &\rightarrow (\forall z \in y) (\exists w) \Psi(w, z, \vec{x}) \quad (\text{by logic}) \\ &\leftrightarrow (\forall z \in y) \Phi(z, \vec{x}). \end{aligned}$$

This provides us with the Σ_1 equivalent

$$(\exists v) (\forall z \in y) (\exists w \in v) \Psi(w, z, \vec{x})$$

to $(\forall z \in y) \Phi(z, \vec{x})$. \square

Now, both in the case of BS and KP, as well as considering these as LST theories, we introduced analogous, set-theoretic notions defined within set theory proper, namely the notions of amenable and admissible sets, respectively. This is to enable us to obtain, within set theory itself, “localised” analogues of some of our later results concerning the logical complexity of the constructible hierarchy, and related notions. By and large, the importance of this will become clear as we progress through Chapter II, but in the meantime, by way of an illustration, we formulate our next result not as a theorem schema for KP, as we did with the previous four lemmas, but rather as a (ZF) theorem about admissible sets. Hopefully, the reader should have no difficulty in reformulating both the statement and the proof of this lemma along the lines of the previous KP-results.

- 11.7 Lemma.** Let M be an admissible set, and let F be a $\Sigma_1(M)$ function over M . If $u \in M$ and $u \subseteq \text{dom}(F)$, then $F \upharpoonright u, F''u \in M$.

Proof. By 10.3, $F \upharpoonright u$ is $\Delta_1(M)$. So by Δ_1 -Comprehension (11.4),

$$w \in M \rightarrow w \cap (F \upharpoonright u) \in M.$$

Now,

$$(\forall x \in u) \exists y [y = F(x)],$$

so by 11.5 (or rather the consequence/analogue of 11.5 for admissible sets) there is a $v \in M$ such that

$$(\forall x \in u) (\exists y \in v) [y = F(x)].$$

Thus $F \upharpoonright u \subseteq v \times u$. But $w = v \times u \in M$, by the Cartesian Product Axiom. Hence

$$F \upharpoonright u = w \cap (F \upharpoonright u) \in M.$$

By Σ_0 -Comprehension now,

$$F'' u = v \cap \{y \mid (\exists x \in u)[(y, x) \in F \upharpoonright u]\} \in M. \quad \square$$

So far we have stated results either as theorem schemas for KP or as theorems within ZF about admissible sets. It is convenient to state the next result as a theorem schema *in terms of classes* (as we often do for ZF). Thus, a Σ_1^{KP} class is a class of the form

$$\{x \mid \Phi(x)\}$$

where Φ is a Σ_1^{KP} formula of LST, etc. And a Σ_1^{KP} function over V is a class of the form

$$\{(y, \vec{x}) \mid \Phi(y, \vec{x})\},$$

such that Φ is a Σ_1 formula of LST and

$$\text{KP} \vdash \forall \vec{x} [\exists y \Phi(y, \vec{x}) \rightarrow \exists! y \Phi(y, \vec{x})].$$

11.8 Lemma (The Recursion Theorem). *Let G be a total, $(n + 2)$ -ary, Σ_1^{KP} function over V . Then there is a total, $(n + 1)$ -ary, Σ_1^{KP} function, F , over V such that:*

$$\text{KP} \vdash F(y, \vec{x}) = G(y, \vec{x}, (F(z, \vec{x}) \mid z \in y)).$$

Proof. Let $\Phi(\sigma, \vec{x})$ be the LST formula

$$\begin{aligned} & [“\sigma \text{ is a function}”] \wedge [“\text{dom}(\sigma) \text{ is transitive}”] \\ & \wedge [(\forall y \in \text{dom}(\sigma))(\sigma(y) = G(y, \vec{x}, \sigma \upharpoonright y))]. \end{aligned}$$

Since G is total, by 10.3, G is in fact a Δ_1^{KP} class. Hence Φ is Δ_1^{KP} . Thus $\Psi(z, y, \vec{x})$ is a Σ_1^{KP} formula, where

$$\Psi(z, y, \vec{x}) = (\exists \sigma) [\Phi(\sigma, \vec{x}) \wedge \sigma(y) = z].$$

Claim 1. $\text{KP} \vdash (\forall \vec{x}, y)(\exists z) \Psi(z, y, \vec{x})$.

Proof of claim: Argue in KP. Suppose otherwise. Pick \vec{x}, y so that

$$\neg (\exists z) \Psi(z, y, \vec{x}).$$

By the Axiom of Foundation, we can ensure that y is chosen here so that

$$(\forall y' \in y)(\exists z) \Psi(z, y, \vec{x}).$$

By 11.5, we can find a set v such that

$$(\forall y' \in y)(\exists \sigma \in v)(y' \in \text{dom}(\sigma) \wedge \Phi(\sigma, \vec{x})).$$

By Δ_1 -Comprehension, set

$$w = v \cap \{\sigma \mid \Phi(\sigma, \vec{x})\}.$$

Let $\varrho = \bigcup w$. Then ϱ is a function. To see this, it clearly suffices to show that if $z \in \text{dom}(\sigma_1) \cap \text{dom}(\sigma_2)$, where $\Phi(\sigma_1, \vec{x})$ and $\Phi(\sigma_2, \vec{x})$, for $\sigma_1, \sigma_2 \in v$, then $\sigma_1(z) = \sigma_2(z)$. But this follows from the nature of Φ by \in -induction: if $\sigma_1(z') = \sigma_2(z')$ for all $z' \in z$, then $\sigma_1 \upharpoonright z = \sigma_2 \upharpoonright z$, and therefore

$$\sigma_1(z) = G(z, \vec{x}, \sigma_1 \upharpoonright z) = G(z, \vec{x}, \sigma_2 \upharpoonright z) = \sigma_2(z).$$

And clearly, $\text{dom}(\varrho)$ is transitive. It is now clear that $\Phi(\varrho, \vec{x})$. Let

$$\tau = \varrho \cup \{(G(y, \vec{x}, \varrho \upharpoonright y), y)\}.$$

Clearly, $\Phi(\tau, \vec{x})$. But

$$\tau(y) = G(y, \vec{x}, \varrho \upharpoonright y).$$

Hence $\Psi(\tau(y), y, \vec{x})$, contrary to the choice of \vec{x}, y . The claim is proved.

Let F be the class

$$\{(z, y, \vec{x}) \mid \Psi(z, y, \vec{x})\}.$$

Claim 2. $\text{KP} \vdash F$ is a function.

Proof of claim: Just as the proof that ϱ was a function in claim 1.

Clearly, F is a required for the lemma. \square

11.9 Corollary. *The function TC (transitive closure) is Σ_1^{KP} (and hence Δ_1^{KP}).* \square

Using 11.9, together with an argument much as in 11.8, we get:

11.10 Lemma (TC-Recursion Theorem). *Let G be a total, $(n + 2)$ -ary, Σ_1^{KP} function over V . Then there is a total, $(n + 1)$ -ary, Σ_1^{KP} function, F , over V such that*

$$\text{KP} \vdash F(y, \vec{x}) = G(y, \vec{x}, (F(z, \vec{x}) \mid z \in \text{TC}(y))). \quad \square$$