1. SUPERSTRUCTURES

In this section and the next we develop the framework for nonstandard analysis that we will use from here on. It is based on viewing mathematical objects as sets, as is usually done in the foundations of mathematics.

In this section we introduce a family of set theoretic structures and a language for the formalization of set theoretic concepts and constructions. Typically we begin an application by choosing an infinite set X of basic entities; for example (as in the previous section) this might be the set of real numbers. We refer to X as a *set of individuals*. We regard the elements of X as *not* being sets. In particular, they do not have elements, yet they are distinct from the empty set. The mathematical objects to which our language will refer are those built over X by elementary set theoretic constructions such as formation of the power set. In a departure from the usual perspective in set theory, we only consider objects that occur at "finite levels" above X; this limitation will be made precise below. This is a pragmatic choice, made for two reasons: first, we are guided by normal mathematical practice, which mainly uses set theoretic objects of "finite height" over some natural base set; second, the nonstandard interpretation of our basic language is easier to comprehend and use if we limit the use of set theory in this way. No loss of mathematical generality results from this limitation; to represent more mathematical objects we can always increase the size of the base set X.

Let X be any set of individuals. For each $n \geq 0$ we define $U_n(X)$ to be the elements of X together with the collection of all sets of "height" or "level" at most n over X. This is made precise by induction on n:

$$\begin{aligned} U_0(X) &= X \\ U_{n+1}(X) &= U_n(X) \cup \mathcal{P}(U_n(X)) \end{aligned}$$

where $\mathcal{P}$ is the power set operation. (That is, $\mathcal{P}(A)$ is the collection of all subsets of A, for any set A.) We also set

$$U(X) = \bigcup_{n \in \mathbb{N}} U_n(X).$$

1.1. **Terminology.** A *superstructure* is a set of the form $U(X)$ where X is an infinite set of individuals; in that situation we refer to $U(X)$ as *the superstructure over* X.

If $a \in U(X)$, we define the *rank* of a to be the least number n for which $a \in U_n(X)$. Everything in $U(X)$ is either a set or an element of X, as is easily proved by induction on rank. Indeed, the elements of X have rank 0 and the sets in $U(X)$ have rank > 0. Moreover, if A is a nonempty set in $U(X)$, then

$$\text{rank}(A) = 1 + \max\{\text{rank}(x) \mid x \in A\}$$

while the empty set has rank 1.

As is usual in set theory, we define the ordered pair $\langle a, b\rangle$ of two objects a, b to be the set $\{\{a\}, \{a, b\}\}$. If A, B are sets, then $A \times B$ is the set of all ordered pairs $\langle a, b\rangle$ for which $a \in A$ and $b \in B$. We treat functions as being sets of ordered pairs. In particular, a function f from A to B is regarded as a subset of $A \times B$.

1.2. **Exercises.** (i) $U_{n+1}(X) = X \cup \mathcal{P}(U_n(X))$.
(ii) $U_n(X) \in U_{n+1}(X)$.
(iii) $U_n(X)$ is transitive; that is, if A is a set in $U_n(X)$ and a is any element of A, then $a \in U_n(X)$.
(iv) If $a, b \in U_n(X)$, then $\{a, b\} \in U_{n+1}(X)$ and $\langle a, b\rangle \in U_{n+2}(X)$.
(v) If A, B are sets in $U_n(X)$, then $A \cup B \in U_n(X)$.
(vi) If A is a set in $U_n(X)$, then every subset of A is also in $U_n(X)$ and $\mathcal{P}(A)$ is in $U_{n+1}(X)$.
(vi) If A, B are sets in $U_n(X)$ then $A \times B \in U_{n+2}(X)$; the set of all functions from A into B is in $U_{n+3}(X)$.
(vii) Let S be any subset of $U(X)$. Then S is an element of $U(X)$ if and only if there exists $n \in \mathbb{N}$ such that $\text{rank}(x) \leq n$ holds for all $x \in S$. If n is such a bound, then $S \in U_{n+1}(X)$.

Constructing superstructures within set theory. Anyone used to formalizing all of mathematics in set theory might find the notion of a "set of individuals" somewhat mysterious. This is because set theorists have come to regard everything as a set, and thus to regard two things as equal if they have the same elements. One way around this is to modify the foundations of set theory so that it allows for individuals (as is done in Mostowski's version of the Zermelo-Fraenkel system). Another way is described next; we consider ourselves as working within a universe of sets V in which the Zermelo-Fraenkel axioms hold. For any X, construct the sets $U_n(X)$ using the same definition as above, interpreted in V:

$$\begin{aligned} U_0(X) &= X \\ U_{n+1}(X) &= U_n(X) \cup \mathcal{P}(U_n(X)) \end{aligned}$$

Note that for some sets X, there may well be some intersection between X and $\mathcal{P}(U_n(X))$ for various n, which would make it problematic to use X as a "set of individuals." However, by choosing X carefully we can avoid this difficulty.

Call X a *set of atoms* if there does not exist any finite sequence $z_0, z_1, \ldots, z_m$ in V $(m \geq 1)$ such that z_0 and z_m are in X and $z_{i-1} \in z_i$ holds for all $i = 1, \ldots, m$. The following facts are easy to prove:

(I) If X is a set of atoms, then $U_{n+1}(X)$ is the disjoint union of X and $\mathcal{P}(U_n(X))$, for each $n \in \mathbb{N}$.

(Proof by induction on n.)

(II) For every cardinal κ, there is a set of atoms X of cardinality κ.

(Proof. Let T be a transitive set of cardinality at least κ, and take X to be any subset of $\mathcal{P}(T)$ that is disjoint from T and has cardinality equal to κ. If $z_0, z_1, \ldots, z_m$ in V $(m \geq 1)$ satisfy $z_m \in X$ and $z_{i-1} \in z_i$ for all $i = 1, \ldots, m$, then $z_0, \ldots, z_{m-1}$ are all in T and hence are not in X. Thus X is a set of atoms.)

Therefore, to model "sets of individuals" in Zermelo-Fraenkel set theory, it suffices to take them to be sets of atoms.

The language of superstructures. We now introduce a first order language L suitable for expressing mathematical statements in superstructures. The nonlogical symbols of L are a binary relation symbol E for membership and a unary relation symbol I to designate the individuals. Given a set X of individuals, we consider $U(X)$ as the L-structure

$$\mathcal{U}(X) = (U(X), \in, X).$$

When writing L-formulas, we will often write $\in$ instead of E.

For most purposes we only need certain formulas from L. These are the *bounded* L-formulas, also called Δ_0*-formulas*, which are defined by induction as follows:

1.3. **Definition.** (i) Every quantifier-free L-formula is a Δ_0-formula.
(ii) Every propositional combination of Δ_0-formulas is a Δ_0-formula.
(iii) If φ is a Δ_0-formula and x, y are distinct variables, then $\forall x\,(x \in y \wedge \varphi)$ and $\exists x\,(x \in y \wedge \varphi)$ are Δ_0-formulas.

For convenience we will abbreviate $\forall x\,(x \in y \wedge \varphi)$ by $\forall x \in y\ \varphi$ and $\exists x\,(x \in y \wedge \varphi)$ by $\exists x \in y\ \varphi$. Note that y is a free variable in these formulas.

A key property of Δ_0-formulas is that they are preserved on passing to transitive substructures. This will be important in the next section, where we discuss nonstandard extensions of superstructures.

1.4. **Definition.** Let $\mathcal{A}$ and $\mathcal{B}$ be L-structures. We say $\mathcal{A}$ is a *transitive* substructure of $\mathcal{B}$ if $\mathcal{A}$ is a substructure of $\mathcal{B}$ and for all a, b in $\mathcal{B}$, if a is in $\mathcal{A}$ and $bE^{\mathcal{B}}a$ holds, then b is also in $\mathcal{A}$.

1.5. **Proposition.** *Let $\mathcal{A}$ and $\mathcal{B}$ be L-structures and suppose that $\mathcal{A}$ is a transitive substructure of $\mathcal{B}$. If $\varphi(x_1, \dots, x_m)$ is a Δ_0-formula and $a_1, \dots, a_m$ are in $\mathcal{A}$, then*

$$\mathcal{B} \models \varphi[a_1, \dots, a_m] \quad \textit{if and only if} \quad \mathcal{A} \models \varphi[a_1, \dots, a_m].$$

Proof. The proof is by induction on Δ_0-formulas. The key point is that whenever one interprets a bounded quantifier $Q\,x \in y$ and y refers to an element of $\mathcal{A}$, then the range of values of x that one needs to consider is the same in $\mathcal{B}$ as in $\mathcal{A}$, by the assumption of transitivity. ☐

Also of importance to us will be the following (trivial) comprehension principle for superstructures.

1.6. **Proposition** (Δ_0-comprehension for superstructures)**.** *Let X be a set of individuals and let $\mathcal{U}(X)$ be the superstructure $(U(X), \in, X)$ based on X. Let x be a set in $U(X)$ and let $y_1, \dots, y_n$ be any elements of $U(X)$. Let $\varphi(u, v_1, \dots, v_n)$ be any Δ_0-formula in L. Then there exists a set $y \subseteq x$ in $U(X)$ such that for all $u \in U(X)$*

$$u \in y \quad \textit{if and only if} \quad \mathcal{U}(X) \models u \in x \wedge \varphi(u, y_1, \dots, y_n).$$

Proof. Immediate. ☐

A very large number of basic set theoretic concepts can be described by Δ_0-formulas. We finish this section by giving some examples that will be of importance later. In each case we are restricting our attention to superstructures $(U(X), \in, X)$ and the variables $x, y, z, \ldots$ are taken to vary over elements of $U(X)$.

(1) "x, y are sets and x is a subset of y" is expressed by the conjunction of $\neg I(x) \wedge \neg I(y)$ and

$$\forall z \in x (z \in y).$$

(2) "y is the union of the sets in x" is expressed by the conjunction of $\neg I(y)$ and

$$\forall w \in y \; \exists z \in x \; (w \in z) \wedge \forall z \in x \; \forall w \in z \; (w \in y).$$

(3) "x, y, z are sets and z is the union of x and y" is expressed by the conjunction of $\neg I(x) \wedge \neg I(y) \wedge \neg I(z)$ and

$$\forall w \in z \; (w \in x \vee w \in y) \wedge \forall w \in x \; (w \in z) \wedge \forall w \in y \; (w \in z).$$

(4) "x, y, z are sets and z is the intersection of x and y" is expressed by the conjunction of $\neg I(x) \wedge \neg I(y) \wedge \neg I(z)$ and

$$\forall w \in z \; (w \in x \wedge w \in y) \wedge \forall w \in x \; (w \in y \rightarrow w \in z).$$

(5) " x, y, z are sets and $z = x \setminus y$" is expressed by the conjunction of $\neg I(x) \wedge \neg I(y) \wedge \neg I(z)$ and

$$\forall w \in z \; (w \in x \wedge w \notin y) \wedge \forall w \in x \; (w \notin y \rightarrow w \in z).$$

(6) "z is the set whose elements are exactly x, y" is expressed by

$$x \in z \wedge y \in z \wedge \forall w \in z \; (w = x \vee w = y).$$

(7) "z is the ordered pair $\langle x, y \rangle$" is expressed by

$$\exists u \in z \; \exists v \in z \; (z = \{u, v\} \wedge u = \{x, x\} \wedge v = \{x, y\}).$$

Note that we are here using the Δ_0-formula in item (6).

(8) "z is an ordered pair and x is the first coordinate of z" is expressed by

$$\exists u \in z \; \exists y \in u \; (z = \langle x, y \rangle).$$

(9) "z is an ordered pair and y is the second coordinate of z" is expressed by

$$\exists u \in z \; \exists x \in u \; (z = \langle x, y \rangle).$$

(10) "x, y are sets and z is the cartesian product of x and y" is expressed by the conjunction of $\neg I(x) \wedge \neg I(y)$ and

$$\forall w \in z\ \exists u \in x\ \exists v \in y\ (w = \langle u, v\rangle) \wedge \forall u \in x\ \forall v \in y\ \exists w \in z\ (w = \langle u, v\rangle).$$

(11) "x is a binary relation" is expressed by

$$\neg I(x) \wedge \forall y \in x\ \exists u \in y\ \exists v \in u\ \exists w \in u\ (y = \langle v, w\rangle).$$

A *binary relation* is a set of ordered pairs.

(12) "x is a binary relation and y is the domain of x" is expressed by the conjunction of item (9) and

$$\forall u \in y\ \exists v \in x\ \exists w \in v\ \exists z \in w\ (v = \langle u, z\rangle) \wedge$$
$$\forall u \in x\ \forall v \in u\ \forall w \in v\ \forall z \in v\ \big(u = \langle w, z\rangle \to w \in y\big).$$

The *domain* of a binary relation is the set of first coordinates of the ordered pairs it contains.

(13) "x is a binary relation and y is the range of x" is expressed by the conjunction of item (9) and

$$\forall u \in y\ \exists v \in x\ \exists w \in v\ \exists z \in w\ (v = \langle z, u\rangle) \wedge$$
$$\forall u \in x\ \forall v \in u\ \forall w \in v\ \forall z \in v\ \big(u = \langle z, w\rangle \to w \in y\big).$$

The *range* of a binary relation is the set of second coordinates of the ordered pairs it contains.

(14) "x is a function" is expressed by the conjunction of item (9) and

$$\forall p \in x\ \forall q \in x\ \forall u \in p\ \forall w \in q\ \forall a \in u\ \forall b \in u\ \forall c \in w$$
$$\big((p = \langle a, b\rangle \wedge q = \langle a, c\rangle) \to b = c\big).$$

(15) "x has rank 0" is expressed by $I(x)$.

(16) "x has rank ≤ 1" is expressed by $\forall x_1 \in x\ (I(x_1))$.

(17) For each $n \geq 2$, the condition "x has rank $\leq n$" is expressed by

$$\forall x_1 \in x\ \forall x_2 \in x_1\ \ldots \forall x_n \in x_{n-1}\ (I(x_n)).$$

(18) "x is a transitive set" is expressed by the conjunction of $\neg I(x)$ and

$$\forall u \in x\ \forall v \in u\ (v \in x).$$