POLSKA AKADEMIA NAUK, INSTYTUT MATEMATYCZNY

DISSERTATIONES MATHEMATICAE
(ROZPRAWY MATEMATYCZNE)

CVI

ARNOLD OBERSCHELP

Set theory over classes

WARSZAWA 1973
PAŃSTWOWE WYDAWNICTWO NAUKOWE

PRINTED IN POLAND

WROCŁAWSKA DRUKARNIA NAUKOWA

CONTENTS

1. Introduction

In this paper we discuss several systems dealing with sets and classes. The main point of the paper is that there is intuitively a difference between sets and classes in general and that proper classes can be elements of sets and act in much the same way as Urelements. In section 2 we give a general discussion of how we look at sets and classes. In section 3 we give a description of the language and underlying logic. A characteristic feature is that the terms need not denote objects of the individual domain. The axiomatic treatment of ordered pairs allows to introduce the elementary theory of relations and functions without any class theoretical existence assumptions. In section 4 we describe the system **ZF + Cls**, the *general set theory over classes*. The question what classes there are is left open in that system. We only require that all sets of classes exist. The system serves as a common starting point for extensions coming from **ZF + Cls** by adding comprehension principles. In section 5 we present the system **ZF +M-Comp** which is almost the same as the system $\mathbf{G}^*$ of [14]. The comprehension principle of this system which yields a lot of proper classes is got from the naive (and contradictory) comprehension schema by a certain relativization process which relativizes to the sets all variables standing in a condition so to speak both for elements and for classes. The investigation of **ZF +M-Comp** is continued in section 6.

In section 7 we present the system **ZF + NF** which combines set theory and stratification and comes from **ZF + Cls** by adding the comprehension principle of Quine's system **NF** ("New foundations") [20], viz. the stratification schema. The system **ZF + NF** is a set theory in which functor categories (e. g. the category of all categories) exist in a natural way. The system **ZF + M-Comp** is too weak for this.

Systems **ZF + M-Comp** and **ZF + NF** are extensions of **ZF + Cls** in quite different directions, each having some advantages and some drawbacks. Of course, one would like to have a supersystem combining the desirable features of both systems, e. g. a common extension. Unfortunately, **ZF + M-Comp** and **ZF + NF** are incompatible. So, in trying to find a supersystem, one has to weaken the one or the other system or introduce two different sorts of classes.

In section 9 we weaken **ZF + M-Comp** and strengthen **ZF + NF**. We define the notion of M-*stratification* and the system **ZF + M-Strat.**

The notion of M-stratification is a generalization of the notion of stratification. Therefore **ZF+M-Strat** is an extension of **ZF+NF**, but not of **ZF+M-Comp**. However, we hope that **ZF+M-Strat** has got enough good properties to be interesting and attractive.

2. General view on sets and classes

We maintain that a reasonable difference can be made between the intuitive notions of sets and classes. The notion of class is related to abstraction, the notion of set is understood as involving a stepwise construction by certain set formation processes.

Suppose that there are certain objects, also called individuals, about which we speak in a first order language using variables $x, y, z, \ldots$ to range over the objects. Then any formula $\varphi(x)$ determines the class

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

i. e. it holds

$$\forall x\big(x \in \{x \mid \varphi(x)\} \leftrightarrow \varphi(x)\big).$$

In most theories there are more classes than abstractions. Therefore abstraction gives only a partial explanation of classes. The distinctive feature of classes is rather that they are uniquely determined by their elements, that they obey the extensionality principle.

The notion of class is symmetric with respect to small and large. Formulas φ and $\neg\varphi$ both define classes. Some special classes are:

$$\emptyset =_{\mathrm{Df}} \{x \mid x \neq x\} \qquad \text{(the empty class)},$$

$$\mathcal{V} =_{\mathrm{Df}} \{x \mid x = x\}\,(^1) \qquad \text{(the universal class)},$$

$$\mathrm{Ru} =_{\mathrm{Df}} \{x \mid x \notin x\} \qquad \text{(the Russell class)}.$$

The universal class contains any object of the theory: It holds

$$\forall x \; x \in \mathcal{V}.$$

The delicate point is what classes are objects of the theory. It is well known that it is impossible that *all* classes over the objects are objects too, e. g. the Russell class is not an object of the theory, i. e. (²) can not belong to the possible values of the variables.

(¹) The symbol $\mathcal{V}$ is adopted from Quine, see [25].

(²) This is "Quine's criterion", see [23]: To be (for a theory) is to be a possible value of the bound variables (in order that the theorems of the theory come out true).

Adopting a term of Quine [25], we call classes *real*, if they are objects of the theory. X is real means: $\exists x\, x = X$, equivalently: $X \in \mathfrak{V}$. The other classes are called *virtual*(³). Thus Ru is virtual, $\mathrm{Ru} \notin \mathfrak{V}$. Since it is not possible to have all classes over the objects to be real, the question arises which classes should be considered to be real. There are satisfactory answers in this respect concerning sets (see below). But beyond the sets there seems to be no intuition at all(⁴). We try to formulate some desirable features which a good choice of real classes should have. In lack of intuitive hints, we only have formalistic and pragmatic criteria. One wants to have as many classes as possible to be real. The real classes should be closed under the Boolean operation (meet, union, complement) and under the operations of the general theory of relations and functions (domain, range, inverse, composition, direct product etc.). Finally for formulas yielding real classes (via abstraction) there should be a simple syntactic criterion. There should be a comprehension principle

$$\textbf{Comp}\colon\ \{x \mid \varphi\} \in \mathfrak{V} \text{ if } \ldots,$$

where "..." indicates a syntactical condition on φ excluding contradictions. These requirements do not uniquely determine a theory. But if we look for a class theory along these lines(⁵), then it seems fair to say that the notion of class is intuitively quite different from the notion of set which we will discuss now.

The notion of set is in set theory understood in quite another way. It is true that Cantor's distinction of "konsistente Vielheiten" and "inkonsistente Vielheiten" corresponds more to the difference of real and virtual classes, and that he identified sets with the former. But fortunately, as we think, a more restricted understanding of sets has come into usage, and we suggest to understand sets in this "Zermeloan" way. Under this interpretation any set is a real class. But it is by no means clear that any abstraction yielding a real class will give a set. Sets are got from the objects at hand by the well-known set formation processes (pairs, unions, powers, subsets, images) described by the Zermelo–Fraenkel axioms. The sets are well-founded relative to the objects one is starting with. Sets are

---

(³) The difference between real and virtual classes is not an absolute one but relative to the theory. Classes virtual for one theory may be real for an extension, but new virtual classes (e. g. a new Russell class) come up in that extension.

(⁴) The antinomies of set theory are rather antinomies of *class* theory. They come up by assuming to many classes to be real. As Gödel has pointed out [8], the notion of set has never led to any contradiction and (empirically) proved to be quite sound.

(⁵) A typical class theory of this type is the theory **NF** ("New Foundations") of Quine [20]. But we think that the notion of set does not occur in **NF**. The same applies to the theory **ML** ("Mathematical Logic") of Quine [21].

asymmetric with respect to small and large. If x is a set, then any smaller collection is a set too.

We know that it is impossible that all classes of objects are objects again. But there is no difficulty in assuming that all sets of objects are objects again. Therefore we consider only theories in which *any* set of objects is an object again(6). By this we decide against ultimate classes (see below).

The elements of classes and sets need not be classes or sets. They may be urelements, i. e. objects without elements but different from the empty class. Actually, often one thinks of some urelements "to start with". However, it has turned out that for mathematics nothing but classes is needed. Therefore we later simplify our theory by leaving outside urelements and classes over urelements (7). We restrict ourselves to *pure* classes. That means that we adopt the extensionality axiom:

$$\textbf{Ext}:\ \forall x \forall y \big(x = y \leftrightarrow \forall z (z \in x \leftrightarrow z \in y)\big).$$

Thus in our theory any element of a class is a class again. But it need not be that any element of a set or class is a set again. It is true that an element of a set or class must be a well-defined object, i. e. real. But it must not be a set or an urelement. *An element may be a real class that is not a set, a proper class.* Actually, this is the main point of the present paper.

In the usual systems of set theory (without urelements) any element is a set. If the predicate "M" means "is a set", then in these systems we have:

$$(*) \qquad x \in \mathrm{M} \leftrightarrow \exists y\ x \in y.$$

But this is only correct because these systems deal with a special sort of sets which we will call *hereditary sets.* A set x is a hereditary set if all elements of x are sets, all elements of elements of x are sets etc. The class of hereditary sets will be called V(8):

$$\mathrm{V} = \textit{class of hereditary sets}.$$

It is true that V contains the most important sets and that for most purposes a theory dealing only with V is sufficient and appropriate. V contains the ordinals and cardinals (in the von Neumann sense), V con-

(6) Another possibility would be to take just the hereditary sets (see below) to be real. Of this kind are the well-known systems **ZF**, **NBC**, and **NQ** (i. e. $\mathbf{ZF}^{(2)}$). There are also systems like Ackermann's, which have more real sets than just the hereditary sets but which only call these "sets".

(7) It is possible to modify our theories in such a way that urelements are admitted.

(8) The symbol "V" is used for the class of sets in some systems of set theory which from our point of view deal with exactly the hereditary sets.

tains ω and $P\omega$. Continuum Hypothesis CH and General Continuum Hypothesis GCH are statements about V. However, we think that the general notion of set — even if we exclude urelements — is broader than the notion of hereditary set. If x is a well-defined object of the theory (i. e. a real class), then certainly the singleton $\{x\}$ being a "small" class, is intuitively a set, even if x is a proper class. In general therefore (*) fails to characterize the sets. (*) is only correct if among the real classes there are only hereditary sets and only classes of (hereditary) sets. Therefore we are led to adopt "M" as a new primitive. The ϵ-language (without "M") is suitable to deal only with classes and will treat sets inadequately except in special situations. In the ϵ-M-language (with both primitives) we can study sets and classes in more general situations (9).

Perhaps one might think that the admission of proper classes as elements of sets will quickly lead into contradictions. This is not at all the case. Moreover, we think that sets over proper classes are in perfect accordance with the intuitive ("Zermeloan") picture of sets used in mathematics. It is true that we cannot adopt the **ZF**-axioms of the theory of hereditary sets literally. This is because they are usually formulated with tacit use of the fact that one is only speaking about V. But we can reformulate the axioms in an appropriate way without changing their intuitive content. This reformulation has been given in [14] and will be repeated here. The formulation is similar to formulations of axioms for set theory over urelements (10). The only difference is that our non-sets have elements and are proper classes and that sets are not defined but given by a primitive notion. The question what classes there are beyond the sets, i. e. what real proper classes there are, is left open in the *general* set theory over classes. For more specific systems we will answer this question later. Before presenting the set theory over classes, we give in the next section a description of the language and underlying logic. The language will contain connectives, quantifiers, and the equality symbol. For convenience we also add the descriptive operator. Our language is a class theoretical language, therefore we have the primitive ϵ and the classifier $\{\ldots|\ldots\}$. Of course, we also have the primitive M. We furthermore add the primitive notion $\langle\ldots,\ldots\rangle$ (a binary operation symbol) for the ordered pairs.

It is true that we later can define ordered pairs and prove the axiom of ordered pairs. However, there are some arguments for an axiomatic treatment of ordered pairs.

(9) Quine believes very strongly in elegance and in economy of the number of primitives. He says (with respect to Ackermann's set theory [1]) that it is "hard to accept the inelegance of the added primitive term $\mathfrak{M}$" (see [25], p. 322). However, we think that classes and sets *are* different and that two primitives are adequate. Moreover, in what follows we will propose a third primitive.

(10) See [27], [30] e. g.

We think that the ordered pair $\langle x, y\rangle$ *is not* (in an absolute way) the same as $\{\{x\}, \{x, y\}\}$, but that this is rather the realization of the intuitive notion of ordered pairs in some systems of set theory or class theory. In other systems there may be other definitions of ordered pairs more convenient. Moreover, any definition seems to involve some existence assumptions about sets or classes. By our choice we avoid *any* commitment and can treat the elementary theory of relations and functions on the "logical level", i. e. without any specific classtheoretical and settheoretical assumptions. A purist in the number of primitives may eliminate the primitive notion later. But the elimination is different in different systems. This is another reason for our axiomatic treatment of ordered pairs.

In our language we can introduce already a lot of terminology concerning classes, relations, and functions and many elementary theorems can be proved. Since it thus looks already like a system, we call it the elementary theory of classes, relations, and functions, and designate it by **CRF**.

3. The elementary theory of classes, relations, and functions

The language has the *variables* $v_0, v_1, v_2, \ldots$ and the *symbols* $\neg, \wedge, \vee, \forall, \exists, =, \epsilon, \iota, \langle\ ,\ \rangle, \{\ ,\ |\ ,\ \}$, and M. We define *general terms* and *formulas.*

(1) *Any variable and* M *is a general term. If X, Y are general terms, x is a variable and φ is a formula, then $\ulcorner\langle X, Y\rangle\urcorner$, $\ulcorner\iota x\varphi\urcorner$, $\ulcorner\{x\,|\,\varphi\}\urcorner$ are general terms (x being bound in the last two terms).*

(2) *If X, Y are general terms, x is a variable and φ, ψ are formulas, then $\ulcorner X = Y\urcorner$, $\ulcorner X \epsilon Y\urcorner$, $\ulcorner\neg\varphi\urcorner$, $\ulcorner(\varphi\wedge\psi)\urcorner$, $\ulcorner(\varphi\vee\psi)\urcorner$, $\ulcorner\exists x\varphi\urcorner$, $\ulcorner\forall x\varphi\urcorner$ are formulas (x being bound in the last two formulas).*

We use $x, y, z, \ldots$ as syntactical variables for variables; φ, ψ for formulas; X, Y, ... (but also $R, Q, f, g, \ldots$) for general terms. Quasiquotation corners are used in the manner of Quine [21].

$\ulcorner\langle X, Y\rangle\urcorner$ is a *pair term*; $\ulcorner\iota x\varphi\urcorner$ is a *description term*; $\ulcorner\{x\,|\,\varphi\}\urcorner$ is an *abstraction term.* Formulas $\ulcorner X = Y\urcorner$, $\ulcorner X\epsilon Y\urcorner$ are *quasiatomic.*

We use the well-known abbreviations (e. g. $\rightarrow$, $\leftrightarrow$, $\overset{1}{\exists}$ (there is exactly one)) and conventions (e. g. concerning brackets) and notions like free and bound occurrences (of variables and general terms). If φ is a formula, than φ^x_X denotes the result of substituting X for x in φ (with the usual precautions, rewriting bound variables if necessary).

If Σ is a set of formulas, then Σ^x_X is the set of φ^x_X for $\varphi \in \Sigma$. Similarily Y^x_X denotes substitution in a general term Y. If $\varphi(X)$ is a formula (with an indicated occurrence of X), then $\varphi(Y)$ is the result of replacing that occurrence by Y. Our general terms may be virtual and not necessarily denote objects. Therefore we have to modify the logic a little bit.

Let us assume that we have a natural deduction type calculus to derive sequents $\Sigma \vdash \varphi$, where Σ is a finite set of formulas. The structural rules, the rules for connectives and quantifiers are as usual. The substitution rule is to be modified as follows:

(3) **Subst.**: *From* $\Sigma \vdash \varphi$ *infer*: $\Sigma^x_X, \ulcorner \exists z\ z = X \urcorner \vdash \varphi^x_X$ (11).

We call a term X for which we can prove $\vdash \ulcorner \exists z\ z = X \urcorner$ an *individual term*. For individual terms the reality premis can be detached. Therefore individual terms can be handled like terms of ordinary first order logic. Concerning identity we have the identity rule and the replacement rule:

(4) **I**: $\vdash \ulcorner X = X \urcorner$ *for any general term* X.
Rpl: $\varphi(X), \ulcorner X = Y \urcorner \vdash \varphi(Y)$ *if the indicated occurrences of* X *and* Y *are free.*

More general replacement rules are then derived as usual. Concerning the descriptive operator, we have the rules for proper description and improper description:

(5) **PD**: $\ulcorner \overset{1}{\exists} x\varphi \urcorner, \varphi \vdash \ulcorner \iota x\varphi = x \urcorner$.
ID: $\ulcorner \neg \overset{1}{\exists} x\varphi \urcorner \vdash \ulcorner \iota x\varphi = \perp \urcorner$.

In the last rule we have already used an abbreviation which is introduced now

(6) $\mathcal{V}$ *is* $\ulcorner \{v_0 | v_0 = v_0\} \urcorner$,
$\emptyset$ *is* $\ulcorner \{v_0 | v_0 \neq v_0\} \urcorner$,
$\perp$ *is* $\ulcorner \iota v_0 v_0 \neq v_0 \urcorner$ (12),
$\ulcorner \mathrm{Cls}(X) \urcorner$ *is* $\ulcorner X = \{z | z \in X\} \urcorner$.

---

(11) Here z is a variable new to the context, say the "next" such variable. Similar requirements apply to other definitions and theorems.

(12) We call $\perp$ the "*phantom object*". It serves as a value when no natural value is at hand, but the value of $\perp$ itself is not determined. One can pick an arbitrary object and identify it with $\perp$.

We have the following rules for classes (x not free in X):

(7) $\mathbf{Cl}_1$: $\vdash \ulcorner X = \{x|\varphi\} \leftrightarrow \mathrm{Cls}(X) \wedge \forall x(x \in X \leftrightarrow \varphi)\urcorner$,
$\mathbf{Cl}_2$: $\ulcorner X \in Y\urcorner \vdash \ulcorner\exists x\, x = X \wedge \mathrm{Cls}(Y)\urcorner$.

Then we can prove the following theorems:

(8) $\vdash \ulcorner\mathrm{Cls}(\{x|\varphi\})\urcorner$,
$\vdash \ulcorner\forall x(x \in \{x|\varphi\} \leftrightarrow \varphi)\urcorner$,
$\ulcorner\mathrm{Cls}(X)\urcorner, \ulcorner\mathrm{Cls}(Y)\urcorner, \ulcorner\forall z(z \in X \leftrightarrow z \in Y)\urcorner \vdash \ulcorner X = Y\urcorner$,
$\vdash \ulcorner\mathrm{Cls}(X) \leftrightarrow X = \emptyset \vee \exists z\, z \in X\urcorner$,
$\vdash \forall x\, x \in \mathcal{V}\urcorner$,
$\vdash \ulcorner\neg\exists x\, x \in \emptyset\urcorner$,
$\vdash \ulcorner\exists z\, z = X \leftrightarrow X \in \mathcal{V}\urcorner$,
$\ulcorner X \in \mathcal{V}\urcorner \vdash \ulcorner\forall x\varphi \rightarrow \varphi^x_X\urcorner$,
$\ulcorner X \in \mathcal{V}\urcorner \vdash \ulcorner\varphi^x_X \rightarrow \exists x\varphi\urcorner$,
$\ulcorner X \in \mathcal{V}\urcorner \vdash \ulcorner X \in \{x|\varphi\} \leftrightarrow \varphi^x_X\urcorner$.

Because of the equivalence for $\ulcorner\exists x\, x = X\urcorner$, we express reality of X usually by $\ulcorner X \in \mathcal{V}\urcorner$.

Next we have rules concerning ordered pairs: the axiom of ordered pairs, a rule concerning the reality of ordered pairs, and a rule concerning pairs with virtual components:

(9) **OP**: $\vdash \ulcorner\langle x, y\rangle = \langle u, v\rangle \leftrightarrow x = u \wedge y = v\urcorner$ ([13]),
ROP: $\vdash \ulcorner\langle x, y\rangle \in \mathcal{V}\urcorner$,
VOP: $\ulcorner X \notin \mathcal{V}\urcorner \vdash \ulcorner\langle X, Y\rangle = \langle Y, X\rangle = \perp\urcorner$ ([14]).

Of course, we need a rule saying that our phantom object $\perp$ is nevertheless real:

(10) **Ph**: $\vdash \ulcorner\perp \in \mathcal{V}\urcorner$.

This phantom axiom is not a deep existence assumption. The logical rules imply anyhow that there are individuals. And (10) does not say more.

([13]) The variables, x, y, u, v shall be different. The same requirement applies without explicit mentioning in analog cases.

([14]) Ordered pairs with virtual components thus have not the characteristic property of ordered pairs. The rule **VOP** has the purpose to give a (purely conventional) value to such pairs. It vould also be possible to drop **VOP** and extend **OP** to a schema with capital letters instead of the small ones. This would have the (rather harmless) effect that one admits virtual non-classes, and the eliminability of virtual terms (described below) would not hold.

Our primitive notion M practically does not play a rôle in this section. We only add a rule saying that M is a class.

(11) **M:** $\vdash \ulcorner \mathrm{Cls}(\mathrm{M}) \urcorner$.

(11) will be provable in set theory.

This ends the description of the system **CRF**, the elementary theory of classes, relations, and functions, which is the underlying language and logic of this paper. By $\vdash \varphi$ we denote that φ is a theorem of **CRF**, by $\Sigma \vdash \varphi$ we denote that φ is derivable in **CRF** from the set of formulas Σ using the rules given above.

Using **Ph** (for the improper cases) one shows:

(12) $\vdash \ulcorner \iota x\varphi, \langle X, Y\rangle \in \mathcal{V} \urcorner$.

Therefore description terms and pair terms are individual terms.

By the admission of general terms which may be real or virtual our language becomes very flexible. But in principle the apparatus of virtual terms can be dispensed with and our system reduced to a system in ordinary first order logic. Our next theorem brings a reduction in that direction.

Let us call a formula *simple* iff the only quasiatomic formulas are of the following types:
$\ulcorner x = y \urcorner$, $\ulcorner x \in y \urcorner$, $\ulcorner x \in \mathrm{M} \urcorner$, $\ulcorner z = \iota x\varphi \urcorner$, $\ulcorner z = \langle x, y\rangle \urcorner$, $\ulcorner z \in \{x \mid \varphi\} \urcorner$, $\ulcorner z = \{x \mid \varphi\} \urcorner$ (where z does not occur in $\ulcorner \iota x\varphi \urcorner$, $\ulcorner \langle x, y\rangle \urcorner$, $\ulcorner \{x, \varphi\} \urcorner$ resp.).

(13) *Any formula is equivalent to a simple formula.*

The proof uses several equivalences which we will not derive here.

(14) $\vdash \ulcorner X = x \leftrightarrow x = X \urcorner$,

$\vdash \ulcorner X = \iota x\varphi \leftrightarrow \exists z(z = X \wedge z = \iota x\varphi) \urcorner$,

$\vdash \ulcorner X = \langle Y, Z\rangle \leftrightarrow \exists x(x = X \wedge x = \langle Y, Z\rangle)$,

$\vdash \ulcorner X = \mathrm{M} \leftrightarrow X = \{x \mid x \in \mathrm{M}\} \urcorner$,

$\vdash \ulcorner X = \{x \mid \varphi\} \leftrightarrow (X = \emptyset \vee \exists y\, y \in X) \wedge \forall y(y \in X \leftrightarrow y \in \{x \mid \varphi\}) \urcorner$,

$\vdash \ulcorner X = \emptyset \leftrightarrow \exists x(x = X \wedge x = \emptyset) \vee (\neg \exists x\, x = X \wedge \forall y\, y \notin X) \urcorner$,

$\vdash \ulcorner X \in Y \leftrightarrow \exists x(x = X \wedge x \in Y) \urcorner$.

Here and below it is understood that the variables that come in on the right-hand side are "new" to the context.

By these equivalences we get a reduction to quasiatomic formulas of the types: $\ulcorner z = X \urcorner$, $\ulcorner z \in X \urcorner$ with z not occurring in X (except X is z).

The types $\ulcorner z = x \urcorner$, $\ulcorner z = \iota x\varphi \urcorner$, $\ulcorner z = \{x|\varphi\} \urcorner$ are allowed in simple formulas. The type $\ulcorner x = \mathrm{M} \urcorner$ has already disappeared. The next equivalence brings a reduction to shorter formulas of type $\ulcorner z = X \urcorner$.

$$(15) \qquad \vdash \ulcorner z = \langle X, Y\rangle \leftrightarrow \exists x \exists y(x = X \wedge y = Y \wedge z = \langle x, y\rangle) \vee (\neg \exists x\, x = X \wedge z = \perp) \vee (\neg \exists y\, y = Y \wedge z = \perp)\urcorner.$$

So we are done with formulas of type $\ulcorner z = X\urcorner$. Next we consider formulas of type $\ulcorner z \in X\urcorner$. The types $\ulcorner z \in x\urcorner$, $\ulcorner z \in \mathrm{M}\urcorner$, $\ulcorner z \in \{x|\varphi\}\urcorner$ are allowed in simple formulas. The next equivalences eliminate the other cases:

$$(16) \qquad \vdash \ulcorner z \in \iota x\varphi \leftrightarrow \exists y(y = \iota x\varphi \wedge z \in y)\urcorner,$$
$$\vdash \ulcorner z \in \langle X, Y\rangle \leftrightarrow \exists y(y = \langle X, Y\rangle \wedge z \in y)\urcorner.$$

A simple formula is not quite an ordinary first order formula. However, using the following equivalences:

$$(17) \qquad \vdash \ulcorner z = \{x|\varphi\} \leftrightarrow (z = \emptyset \vee \exists y\, y \in z) \wedge \forall y(y \in z \leftrightarrow y \in \{x|\varphi\})\urcorner,$$
$$\vdash \ulcorner z \in \{x|\varphi\} \leftrightarrow \varphi_z^x\urcorner,$$

we can eliminate quasiatomic formulas of types $\ulcorner z = \{x|\varphi\}\urcorner$, $\ulcorner z \in \{x|\varphi\}\urcorner$ and only retain $\ulcorner z = \emptyset\urcorner$. By further assumptions (e. g. extensionality or introduction of a primitive constant), we could get rid of $\ulcorner z = \{v_0|v_0 \neq v_0\}\urcorner$ too. Similarily one can use the following equivalence:

$$(18) \qquad \vdash \ulcorner z = \iota x\varphi \leftrightarrow \big(\overset{1}{\exists} x\varphi \wedge \forall x(x = z \leftrightarrow \varphi)\big) \vee (\neg \overset{1}{\exists} x\varphi \wedge z = \perp)\urcorner$$

to eliminate description terms and only retain $\ulcorner z = \perp\urcorner$. By introducing a primitive constant, we could get rid of $\ulcorner z = \iota v_0\, v_0 \neq v_0\urcorner$ too.

It would in principle be possible to start with a first order system and introduce the quasiatomic formulas of our flexible language as abbreviations by context definitions corresponding to the equivalences above.

The additional rules would become eliminable rules of ordinary first order logic. However, this would involve a lot of simple but tedious work. Since we want to have the flexible language anyhow, we find it simpler to introduce it as the basic language with a logic of its own outright.

In our language we can define almost all the concepts occurring in set theory and class theory. First we introduce some terminology of the algebra of classes:

(19) $\ulcorner X \cap Y\urcorner$ *is* $\ulcorner\{u|u \in X \wedge u \in Y\}\urcorner$,
$\ulcorner X \cup Y\urcorner$ *is* $\ulcorner\{u|u \in X \vee u \in Y\}\urcorner$,
$\ulcorner -X\urcorner$ *is* $\ulcorner\{u|u \notin X\}\urcorner$,
$\ulcorner X \setminus Y\urcorner$ *is* $\ulcorner\{u|u \in X \wedge u \notin Y\}\urcorner$,
$\ulcorner X \subseteq Y\urcorner$ *is* $\ulcorner\forall u(u \in X \rightarrow u \in Y)\urcorner$.

All the usual theorems of the algebra of classes are provable.

The next definitions introduce some concepts used in set theory:

(20) $\ulcorner\bigcup X\urcorner$ *is* $\ulcorner\{u\,|\,\exists y(y\in \mathrm{M}\wedge u\in y\in X)\}\urcorner$,
$\ulcorner\bigcap X\urcorner$ *is* $\ulcorner\{u\,|\,\forall y\,(y\in \mathrm{M}\rightarrow(y\in X\rightarrow u\in y))\}\urcorner$.

We call this *set-theoretical union* and *intersection*. $\bigcup X$ thus contains only the elements of *set*-elements of X, non-set-elements are treated as if they were empty. It would also be possible to define a *strong union* and *intersection*:

(21) $\ulcorner\mathbb{U}X\urcorner$ *is* $\ulcorner\{u\,|\,\exists y\; u\in y\in X\}\urcorner$,
$\ulcorner\mathbb{\Pi}X\urcorner$ *is* $\ulcorner\{u\,|\,\forall y(y\in X\rightarrow u\in y)\}\urcorner$.

For $X\subseteq M$ the set theoretical and the strong notions coincide. In set theory we need rather $\bigcup$ than $\mathbb{U}$.

(22) $\ulcorner\mathrm{P}X\urcorner$ *is* $\ulcorner\{u\,|\,u\in \mathrm{M}\wedge u\subseteq X\}\urcorner$.

$\mathrm{P}\,X$ is the class of all *subsets* of X. We can also introduce the notion of *strong powerclass*:

(23) $\ulcorner\mathbb{P}X\urcorner$ *is* $\ulcorner\{u\,|\,\mathrm{Cls}(u)\wedge u\subseteq X\}\urcorner$.

For $X\in\mathrm{M}$ both notions will coincide (by Aussonderungsschema).
It is useful to have the general *abstraction operator*:

(24) *Let X be a general term, φ a formula, and $x_1, \ldots, x_n$ different variables, and z be new. Then* $\ulcorner\{X\,|_{x_1,\ldots,x_n}\varphi\}\urcorner$ *is* $\ulcorner\{z\,|\,\exists x_1\ldots\exists x_n\,(z=X\wedge\varphi)\}\urcorner$.

The variables $x_1, \ldots, x_n$ are often omitted and must be infered from the context.

Relational abstraction is defined by:

(25) $\ulcorner\{x, y\,|\,\varphi\}\urcorner$ *is* $\ulcorner\{\langle x, y\rangle\,|_{x,y}\varphi\}\urcorner$.

The next definitions introduce some terminology involving relations:

(26) $\ulcorner X\times Y\urcorner$ *is* $\ulcorner\{x, y\,|\,x\in X\wedge y\in Y\}\urcorner$,
$\ulcorner\mathrm{Relation}(R)\urcorner$ *is* $\ulcorner R\subseteq \mathcal{V}\times\mathcal{V}\urcorner$,
$\ulcorner XRY\urcorner$ *is* $\ulcorner\langle X, Y\rangle\in R\urcorner$,
$\ulcorner\mathrm{id}_X\urcorner$ *is* $\ulcorner\{x, y\,|\,x=y\in X\}\urcorner$,
$\ulcorner R^{-1}\urcorner$ *is* $\ulcorner\{x, y\,|\,yRx\}\urcorner$,
$\ulcorner R\circ Q\urcorner$ *is* $\ulcorner\{x, y\,|\,\exists z\, xRzQy\}\urcorner$,
$\ulcorner R\restriction X\urcorner$ *is* $\ulcorner\{x, y\,|\,xRy\wedge y\in X\}\urcorner$,
$\ulcorner\mathrm{dom}\; R\urcorner$ *is* $\ulcorner\{x\,|\,\exists y\; yRx\}\urcorner$,
$\ulcorner\mathrm{ran}\, R\urcorner$ *is* $\ulcorner\{y\,|\,\exists x\; yRx\}\urcorner$.

The usual theorems of the general theory of relations hold for these notions.

We mention the following theorems:

(27) $\vdash \ulcorner \mathrm{Relation}(\{x, y \mid \varphi\})\urcorner$,
$\vdash \ulcorner \forall x \forall y (x\{x, y \mid \varphi\} y \leftrightarrow \varphi)\urcorner$,
$\vdash \ulcorner \mathrm{Relation}(R) \wedge \mathrm{Relation}(Q) \wedge \forall x \forall y (xRy \leftrightarrow xQy) \to R = Q\urcorner$.

Our next definitions concern functions:

(28) $\ulcorner \mathrm{Function}(f)\urcorner$ *is* $\ulcorner \mathrm{Relation}(f) \wedge \forall x \forall y_1 \forall y_2 (y_1 f x \wedge y_2 f x \to y_1 = y_2)\urcorner$,
$\ulcorner f \mid X \to Y\urcorner$ *is* $\ulcorner \mathrm{Function}(f) \wedge \mathrm{dom} f = X \wedge \mathrm{ran} f \subseteq Y\urcorner$,
$\ulcorner f(x)\urcorner$ *is* $\ulcorner \iota y\, y f x\urcorner$ (15).

The usual theorems of the general theory of functions hold for these notions.

We mention the extensionality property of functions:

(29) $\vdash \ulcorner \mathrm{Function}(f) \wedge \mathrm{Function}(g) \wedge \mathrm{dom} f = \mathrm{dom}\, g$
$\wedge \forall x \big(x \in \mathrm{dom} f \to f(x) = g(x)\big) \to f = g\urcorner$.

We also want to introduce *functional abstraction*:

(30) $\ulcorner \langle x \mapsto X \mid \varphi\rangle\urcorner$ *is* $\ulcorner \{y, x \mid y = X \wedge \varphi\}\urcorner$,
$\ulcorner \langle x \mapsto X\rangle\urcorner$ *is* $\ulcorner \langle x \mapsto X \mid x \in \mathfrak{V}\rangle\urcorner$.

For the first term one finds also the notation $\ulcorner \langle X |_x \varphi\rangle\urcorner$, the second term is often written as $\ulcorner \lambda x X\urcorner$.

We have the theorem:

(31) $\vdash \ulcorner \forall x (\varphi \to X \in \mathfrak{V}) \to \mathrm{Function}(\langle x \mapsto X \mid \varphi\rangle) \wedge \mathrm{dom}\langle x \mapsto X \mid \varphi\rangle$
$= \{x \mid \varphi\} \wedge \forall x \big(\varphi \to \langle x \mapsto X \mid \varphi\rangle(x) = X\big)\urcorner$.

n-ary relations are just classes of n-tuples, and n-ary functions are functions with an n-ary relation as domain.

We introduce n-tuples by the recursive definition:

(32) $\ulcorner \langle X\rangle\urcorner$ *is* X,
$\ulcorner \langle X_0, X_1, \ldots, X_n\rangle\urcorner$ *is* $\ulcorner \langle X_0, \langle X_1, \ldots, X_n\rangle\rangle\urcorner$.

Then we define:

(33) $\ulcorner R X_1, \ldots, X_n\urcorner$ *is* $\ulcorner \langle X_1, \ldots, X_n\rangle \in R\urcorner$,
$\ulcorner f(X_1, \ldots, X_n)$ *is* $f(\langle X_1, \ldots, X_n\rangle)\urcorner$.

(15) Note that any function "maps" any virtual class onto the phantom object $\perp$.

Other familiar notions like $\ulcorner X^n \urcorner$, $\ulcorner\{x_1, \ldots, x_n | \varphi\}\urcorner$, $\ulcorner\langle x_1, \ldots, x_n \mapsto X\rangle\urcorner$ etc. can be introduced in an obvious way.

$$\begin{array}{lll}
(34) & \ulcorner\{x_1, \ldots, x_n | \varphi\}\urcorner & \textit{is } \ulcorner\{y | \exists x_1 \ldots \exists x_n (y = \langle x_1, \ldots, x_n\rangle \wedge \varphi)\}\urcorner, \\
 & \ulcorner\langle x_1, \ldots, x_n \mapsto X | \varphi\rangle\urcorner & \textit{is } \ulcorner\{y, x_1, \ldots, x_n | y = X \wedge \varphi\}\urcorner, \\
 & \ulcorner\langle x_1, \ldots, x_n \mapsto X\rangle\urcorner & \textit{is } \ulcorner\langle x_1, \ldots, x_n \mapsto X | x_1, \ldots, x_n \in \mathcal{V}\rangle\urcorner, \\
 & \ulcorner \mathrm{pr}_{n,k} \urcorner & \textit{is } \ulcorner\langle v_1, \ldots, v_n \mapsto v_k\rangle\urcorner \ (\textit{for } 1 \leqslant k \leqslant n).
\end{array}$$

The notation in (33) is usual in mathematics. However, one must be careful to avoid ambiguities which could arise since in our flexible language any term may play the rôle of a relation symbol or function symbol or argument symbol. Often the choice of letters "R", "f", "X", indicates what is meant. But a more transparent and systematic method would be desirable. One could think of introducing application symbols, e. g. ϵ_n, $\ni_n$ for relational application, and "$\multimap_n$", "$\multimapinv_n$" for functional application. Thus $\ulcorner X_1 \ldots X_n \epsilon_n X\urcorner$ and $\ulcorner X \ni_n X_1 \ldots X_n\urcorner$ would be a formula and the same as $\ulcorner\langle X_1, \ldots, X_n\rangle \epsilon X\urcorner$, and $\ulcorner X_1 \ldots X_n \multimap_n X\urcorner$ and $\ulcorner X \multimapinv_n X_1 \ldots X_n\urcorner$ would be a term and the same as $\ulcorner \iota z \langle z, X_1, \ldots, X_n\rangle \epsilon X\urcorner$. Then ϵ_1 is the same as ϵ. $\multimapinv_1$ is sometimes written as an accent([10]).

Finally we define the notions of *real classes*, *elements*, *urelements*, *proper classes*, and *ultimate classes*:

$$\begin{array}{ll}
(35) & \mathrm{Cls} \textit{ is } \ulcorner\{v_0 | \mathrm{Cls}(v_0)\}\urcorner, \\
 & \mathrm{EL} \textit{ is } \ulcorner\{v_0 | \exists v_1 v_0 \epsilon v_1\}\urcorner, \\
 & \mathrm{UR} \textit{ is } \ulcorner\mathcal{V} \setminus \mathrm{Cls}\urcorner, \\
 & \mathrm{PC} \textit{ is } \ulcorner\mathrm{Cls} \setminus \mathrm{M}\urcorner, \\
 & \mathrm{UC} \textit{ is } \ulcorner\mathrm{Cls} \setminus \mathrm{EL}\urcorner.
\end{array}$$

Proper classes and ultimate classes are thus real classes. Virtual classes are not even proper classes or ultimate classes.

Our theory **CRF** already looks somewhat like a set theory or class theory. But actually it is nothing but first order logic modelled into a class theoretical language, since we have no assumptions at all about the reality of classes. Our only commitment is that the individual domain is closed under the formation of ordered pairs. But this is a "first-order commitment".

The weakness of our system is shown by the existence of models with exactly one element e, which is also the only ordered pair: $e = \langle e, e\rangle$, and which, of course, must be the phantom object: $e = \perp$. There are two classes $\emptyset$ and $\mathcal{V}$ (which is $\{e\}$). We can put $e \epsilon e$. Then $e = \mathcal{V}$ and $\mathcal{V}$ is

([10]) A symbol like $\multimapinv$ has the advantage of being similar (but not identical) to the composition symbol, the associativity of composition and application reads: $(f \circ g) \multimapinv x = f \multimapinv (g \multimapinv x)$. Moreover, it can be used in reversed form: $x \multimap y$ without ambiguity as to what is the function and what is the argument.

real and $\emptyset$ is virtual. We can also put $e \notin e$, then $\mathfrak{V}$ is virtual. In addition, we can put $e = \emptyset$ (then $\emptyset$ is real), or $e \neq \emptyset$, in the last case e is an urelement and $\mathfrak{V}$, $\emptyset$ are both virtual. But note that finite domains with more than one element are excluded. However, any infinite domain can be admitted, since for an infinite domain there are two-place one-one functions which can simulate ordered pairs.

Our elementary frame is not already a true set theory or class theory. This is only got when we add various assumptions about the reality of sets or classes. But the system **CRF** can serve as a common basis of different set theories and class theories. A comparison of different systems, e. g. as in Quine's book [25] is easily possible on the basis of **CRF**. However, in this paper we are interested only in systems without urelements and containing the general set theory over classes which will be described now.

4. The general set theory over classes

We now give the axioms of the system **ZF+Cls**. Our axioms will not be independent. **AO** follows from **S 4**, **S 4** follows from **AO** and **S 5**.

We first decide to adopt the extensionality axiom:

(36) **Ext**: $\ulcorner x = y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y)\urcorner$ [17].

The next axiom tells us that the empty class is a set.

(37) **AO**: $\ulcorner \emptyset \in \mathrm{M}\urcorner$.

In particular the empty set is a real object. Therefore we can identify the phantom object with the empty set. This is done in the next axiom.

(38) **AO′**: $\ulcorner \perp = \emptyset\urcorner$.

For the purpose of section 3 it was natural to keep $\emptyset$ and $\perp$ independently, e. g. $\emptyset$ could be virtual while $\perp$ is always real. But for set theory the identification in **AO′** is reasonable. But then it would be still simpler to write in (5), (9) $\emptyset$ instead of $\perp$ outright, then (10) and (38) become dispensable. So we agree now that we have made these changes and can forget about **AO′** and the phantom object from now on. **AO** implies $\ulcorner \mathrm{Cls}(\mathrm{M})\urcorner$ and therefore rule (11) is dispensable. Using **Ext** and **AO** one can show:

(39) $\vdash_1 \ulcorner \mathrm{Cls} = \mathfrak{V}\urcorner$.

In particular any set is a class:

(40) $\vdash_1 \ulcorner \mathrm{M} \subseteq \mathrm{Cls}\urcorner$.

[17] The axioms shall be rather the universal closures of the formulas written down.

However, it would easily be possible to admit urelements and set up a system **ZFU+Cls** (general set theory over classes *with urelements admitted*). We only have to drop the extensionality axiom and take (40) as a new axiom instead. No other changes are necessary.

The pair set axiom is often given in the form

(41) $\ulcorner\{x, y\} \in \mathfrak{V} \rightarrow \{x, y\} \in \mathrm{M}\urcorner$.

There are theories in which $\{x, y\}$ is not always real. Any theory with *ultimate classes* ([18]) (i. e. classes that are real but are not elements) is of this kind. In such a theory there are *sets* of objects, e. g. $\{x\}$ for ultimate x, that are virtual and thus no object of the theory ([19]). In such a theory *some* sets and *some* classes are real (often the real sets are just the hereditary sets and the real classes are classes of sets). However, we already decided to have *all* sets and *some* classes to be real (since it is impossible to have all classes as real). Therefore our *pair set axiom* can only be:

(42) **A 1:** $\ulcorner\{x, y\} \in \mathrm{M}\urcorner$.

In adopting **A 1** we decide against ultimate classes. Any real class is an element, even of a set. We have the theorem:

(43) $\vdash_1 \ulcorner\mathrm{UC} = \emptyset\urcorner$.

But $\ulcorner\mathrm{PC} \neq \emptyset\urcorner$ is possible and will hold in later systems.

The union axiom of set theory obviously means that for any set of *sets* the elements of the elements form again a set. Therefore non-set-members do not contribute to the *set*-theoretical union. This means that we have to use $\bigcup$ rather than $\Cup$ in the *union axiom*:

(44) **A 2:** $\ulcorner x \in \mathrm{M} \rightarrow \bigcup x \in \mathrm{M}\urcorner$.

Similar the *power set axiom* is:

(45) **A 3:** $\ulcorner x \in \mathrm{M} \rightarrow \mathrm{P}x \in \mathrm{M}\urcorner$.

Our next axiom is the Aussonderungsaxiom which says that any real subclass of a set is a set:

(46) $\ulcorner x \in \mathrm{M} \rightarrow x \cap y \in \mathrm{M}\urcorner$.

([18]) It is a common prejudice that "proper class" and "ultimate class" mean the same. In this paper we try to argue against this. The term "ultimate class" has been introduced by Quine. However, we do not agree with Quine in our understanding of the term "proper class" and of "set".

([19]) Theories **NBG** and **NQ** (i. e. **ZF**$^{(2)}$) and **ML** have ultimate classes and thus virtual sets.

However, then there might be virtual subclasses of a set which are not subsets ([20]). But it is in accordance with the set intuition that *any* well-defined property on a set will define a subset. This gives an axiom-schema rather than a single axiom.

Therefore the *Aussonderungsschema* is:

(47) **S 4**: $\ulcorner x \in \mathrm{M} \rightarrow x \cap \{u \mid \varphi\} \in \mathrm{M}^\urcorner$.

In Mathematics (for an adequate treatment of induction) we need **S 4** and not (46). The Aussonderungsschema has been generalized by Fraenkel to the *Ersetzungsschema* which says that any image of a set is a set, any class with set-many elements is real and a set.

(48) **S 5**: $\ulcorner x \in \mathrm{M} \wedge (\forall y \in x) \overset{1}{\exists} z \varphi \rightarrow \{z \mid (\exists y \in x)\varphi\} \in \mathrm{M}^\urcorner$.

In this axiomschema formula φ describes a (perhaps virtual) relation which is on the set x a function. **S 4** says that the class of the function values is again a set (and thus real).

It may be noted that if we had at least one pair set, e. g. if $\{\emptyset, \{\emptyset\}\} \in \mathrm{M}$, then **A 1** follows from **S 5**, since $\{x_1, x_2\}$ is the image of $\{\emptyset, \{\emptyset\}\}$ under the function described by the formula: $(y = \emptyset \wedge z = x_1) \vee (y = \{\emptyset\} \wedge z = x_2)$, This means that in theories with ultimate classes (where **A 1** is false). one has to restrict the Ersetzungsschema.

The next axiom is the *Fundierungsaxiom*:

(49) **A 6**: $\ulcorner x \in \mathrm{M} \wedge (\exists y \in x) y \in \mathrm{M} \rightarrow (\exists y \in x)$
$(y \in \mathrm{M} \wedge \neg (\exists z \in x)(z \in \mathrm{M} \wedge z \in y))^\urcorner$.

A 6 says that any set having set elements has also a set element which is minimal with respect to $\in$ among the set elements. A consequence of **A 6** is that there are no descending infinite $\in$-chains of sets. However, infinite descending chains like:

$$\ldots \mathcal{V} \in \mathcal{V} \in \mathcal{V}$$

are of course not excluded. The classes need not be well-founded. But any set is well-founded relative to the non-sets from which it is made up (even if these constituents of set theory are not well-founded themselves). **A 6** guarantees that the sets are just what we get from non-sets by the set axioms (see (72)).

It may be noted that **A 6** is equivalent to a corresponding Fundierungs-schema saying that any class X (virtual or not) such that $X \cap \mathrm{M} \neq \emptyset$ has a set member which is $\in$-minimal among the set members of A.

([20]) There are theories in which any property over sets defines a real class. In these theories (46) and (47) are equivalent. However, this is only possible if the theory deals with "some" sets (e. g. hereditary sets) and not with "all" sets of objects. See (76).

There are two more axioms which we give without further comment. The *infinity axiom* is:

(50) **Inf**: $\ulcorner \exists x\big(x \in \mathrm{M} \wedge \emptyset \in x \wedge (\forall z \in x) z \cup \{z\} \in x\big)\urcorner$.

The *axiom of choice* is:

(51) **AC**: $\ulcorner \forall x\Big(x \in \mathrm{M} \wedge x \subseteq \mathrm{M} \wedge (\forall z \in x) z \neq \emptyset \wedge (\forall z \in x)(\forall y \in x)$
$(z = y \vee z \cap y = \emptyset) \rightarrow \exists w\big(w \in \mathrm{M} \wedge (\forall y \in x)(\exists v \in w) w \cap y = \{v\}\big)\Big)\urcorner$

ZF+Cls is the theory with axioms:

Ext, A 0, A 1, A 2, A 3, S 4, S 5, A 6, Inf, AC.

We write $\Sigma \vdash_1 \varphi$ to indicate that $\Sigma \cup A \vdash \varphi$ for some set A of (closed) formulas, which are axioms of **ZF+Cls** (we did already so in (39), (40), (43)). Thus $\vdash_1 \varphi$ means that φ is a theorem of **ZF+Cls**.

The system **ZFU+Cls** is got from **ZF+Cls** by dropping **Ext** and taking (40) as axiom. The development of the standard body of the theorems of set theory starting from these axioms is possible in the same way as in usual set theory (of hereditary sets or sets over urelements).

We list some theorems:

(52) $\vdash_1 \ulcorner \mathrm{M} \notin \mathrm{M}\urcorner$,
$\vdash_1 \ulcorner \mathcal{V} \notin \mathrm{M}\urcorner$,
$\ulcorner X \in \mathrm{M}\urcorner \vdash_1 \ulcorner -X \notin \mathrm{M}\urcorner$,
$\ulcorner X, Y \in \mathrm{M}\urcorner \vdash_1 \ulcorner X \times Y, X \cap Y, X \cup Y \in \mathrm{M}\urcorner$,
$\ulcorner R, Q \in \mathrm{M}\urcorner \vdash_1 \ulcorner R^{-1}, R \circ Q \in \mathrm{M}\urcorner$,
$\ulcorner \mathrm{Relation}(R)\urcorner \vdash_1 \ulcorner R \in \mathrm{M} \leftrightarrow \mathrm{dom} R, \mathrm{ran} R \in \mathrm{M}\urcorner$,
$\ulcorner \mathrm{Function}(f)\urcorner \vdash_1 \ulcorner f \in \mathrm{M} \leftrightarrow \mathrm{dom} f \in M\urcorner$.

It is well known that the sets $\{\{x\}, \{x, y\}\}$ have all the properties required by $\langle x, y\rangle$. So it would be possible to identify the ordered pair $\langle x, y\rangle$ with $\{\{x\}, \{x, y\}\}$ which we call the Kuratowski pair:

(53) *Possible additional axiom* **OP′**:

$$\ulcorner \langle x, y\rangle = \{\{x\}, \{x, y\}\}\urcorner.$$

This is often done. One can then eliminate the primitive notion and consider **OP′** rather as a definition (of a meta linguistic abbreviation) than as an axiom (in the object language). But we will keep the primitive ordered pairs since we want to have the possibility to redefine them in some other way if necessary. We will review briefly some other pairs we use below. A rather simple pair was given by Schmidt [28]:

(54) *Possible additional axiom* **OP″**:

$$\ulcorner \langle x, y\rangle = \big\{\{\{u\}\} \mid u \in x\big\} \cup \big\{\{\emptyset, \{v\}\} \mid v \in y\big\}\urcorner$$

Another definition was given by Kühnrich [12]:

(55) *Possible additional axiom* $\mathbf{OP}'''$:

$$\ulcorner\langle x, y\rangle = \Big\{\{\{u\}\} \mid u \in x \vee (u \in \mathrm{EL} \wedge u = x)\Big\}$$
$$\cup \Big\{\{\emptyset, \{v\}\} \mid v \in y \vee (v \in \mathrm{EL} \wedge v = y)\Big\}\urcorner.$$

All definitions are in our theory adequate, i. e. the properties ordered pairs should have are fulfilled. (The reality of $\langle x, y\rangle$ under (54), (55) needs a modest portion of class theory beyond set theory.) Under any definition it holds:

(56) $\ulcorner x, y \in \mathrm{M} \to \langle x, y\rangle \in \mathrm{M}\urcorner$
$\ulcorner x, y \in \mathrm{V} \to \langle x, y\rangle \in \mathrm{V}\urcorner$ (adopting $\mathbf{OP}'$, or $\mathbf{OP}''$, or $\mathbf{OP}'''$).

Under (53) it even holds thet *all* pairs are sets:

(57) $\ulcorner\langle x, y\rangle \in \mathrm{M}\urcorner$ (adopting $\mathbf{OP}'$).

In a set theory with ultimate classes the Kuratowski pairs will not work. The Schmidt pairs work for classes but fail for urelements. The Kühnrich pairs will work for urelements and classes (ultimate or not). Therefore these pairs are adequate in the most general situation. If one only has classes (as we have in **ZF+Cls**), then the Schmidt pairs are simpler. If one only has elements (as we have in **ZF+Cls**), then the Kuratowski pairs are still simpler.

There are two notions of transitivity. *Strong transitivity* is:

(58) $\ulcorner\mathrm{Trans}\ X\urcorner$ *is* $\ulcorner\forall x \forall y (x \in y \in X \to x \in X)\urcorner$.

This is equivalent to $\ulcorner\forall y (y \in X \to y \subseteq X)\urcorner$ *and to* $\ulcorner\bigcup X \subseteq X\urcorner$.

Transitivity in the sets is:

(59) $\ulcorner\mathrm{Trans}_M\ X\urcorner$ *is* $\ulcorner\forall x \forall y (y \in \mathrm{M} \wedge x \in y \in X \to x \in X)\urcorner$.

This is equivalent to $\ulcorner\forall y (y \in \mathrm{M} \wedge y \in X \to y \subseteq X)\urcorner$ *and to* $\ulcorner\bigcup X \subseteq X\urcorner$.

In set theory we need rather the last notion. On classes of sets both notions coincide.

(60) $\vdash_1 \ulcorner X \subseteq \mathrm{M} \to (\mathrm{Trans}\, X \leftrightarrow \mathrm{Trans}_M X)\urcorner$.

But actually a transitive class of sets is a class of hereditary sets and thus a subclass of V.

For any set we can define the *transitive closure*

(61) $\ulcorner\mathrm{TC}(X)\urcorner$ *is* $\ulcorner\bigcap\{y \mid y \in \mathrm{M} \wedge \mathrm{Trans}_{\mathrm{M}} y \wedge X \subseteq y\}\urcorner$.

It can be shown that $\mathrm{TC}(X) = X \cup \bigcup X \cup \bigcup\bigcup X \cup \bigcup\bigcup\bigcup X \cup \ldots$ (using recursion over natural numbers and Ersetzungsschema) and that $\mathrm{TC}(X)$ is a set and transitive in the sets:

(62) $\vdash_1 \ulcorner x \in \mathrm{M} \rightarrow x \subseteq \mathrm{TC}(x) \wedge \mathrm{TC}(x) \in \mathrm{M} \wedge \mathrm{Trans}_{\mathrm{M}}\, \mathrm{TC}(x)\urcorner$.

Now we can define the *hereditary sets*:

(63) V *is* $\ulcorner\{v_0 \mid v_0 \in \mathrm{M} \wedge \mathrm{TC}(v_0) \subseteq \mathrm{M}\}\urcorner$.

The *ordinals* are defined as usual:

(64) On *is* $\ulcorner\{v_0 \mid v_0 \in \mathrm{M} \wedge v_0 \subseteq \mathrm{M} \wedge \mathrm{Trans}_{\mathrm{M}} v_0 \wedge (\forall v_1 \in v_0)(\forall v_2 \in v_0)(v_1 \in v_2 \vee v_1 = v_2 \vee v_2 \in v_1)\}\urcorner$.

It turns out that they are hereditary sets:

(65) $\vdash_1 \ulcorner \mathrm{On} \subseteq \mathrm{V}\urcorner$.

Cardinals can be defined as initial ordinals and any set has a cardinal which is in V. One can introduce *transfinite induction and recursion* over the ordinals just as in ordinary set theory. Arguments and values of functions defined by recursions may be any real classes. An example is the *von Neumann hierarchy* over any set of non-sets.

(66) *Let* $z \subseteq \mathfrak{V} \setminus \mathrm{M}$, $z \in \mathrm{M}$ *and define* $\mathrm{V}_\alpha(z)$ *for* $\alpha \in \mathrm{On}$ *by recursion*:

$$\mathrm{V}_0(z) =_{\mathrm{Df}} z,$$
$$\mathrm{V}_{\alpha+1}(z) =_{\mathrm{Df}} \mathrm{P}\mathrm{V}_\alpha(z) \cup z,$$
$$\mathrm{V}_\lambda(z) =_{\mathrm{Df}} \bigcup_{\gamma \in \lambda} \mathrm{V}_\gamma(z) \quad (\textit{for limit ordinal } \lambda),$$
$$\mathrm{V}(z) =_{\mathrm{Df}} \bigcup_{\alpha \in \mathrm{On}} \mathrm{V}_\alpha(z),$$

$\mathrm{V}(z) \setminus z$ are *the sets over* z.

For any set x the *base* of x are the non-sets in $\mathrm{TC}(x)$.

(67) $\ulcorner\mathrm{Base}(X)\urcorner$ *is* $\ulcorner\mathrm{TC}(X) \setminus \mathrm{M}\urcorner$.

Using the Fundierungsaxiom it can be shown that any set is in the von Neumann hierarchy over its base.

(68) $\vdash_1 \ulcorner \forall x (x \in \mathrm{M} \rightarrow x \in \mathrm{V}(\mathrm{Base}(x)))\urcorner$.

This gives the possibility to define a *rank* for any set. Let "$\mu\alpha \ldots$" mean "the least ordinal α such that …"

(69) $\ulcorner\mathrm{rk}(X)\urcorner$ *is* $\ulcorner \mu\alpha\, X \in \mathrm{V}_{\alpha+1}(\mathrm{Base}(X))\urcorner$.

Any set of non-sets has rank zero.

It turns out that the hereditary sets are just the sets with empty base.

(70) $\vdash_1 \ulcorner \mathrm{V} = \mathrm{V}(\emptyset)\urcorner$.

We can also define the von Neumann hierarchy over arbitrary classes of non-sets.

(71) *Let* $X \subseteq \mathfrak{V}\setminus \mathrm{M}, X \notin \mathrm{M}$, *and* a *be an ordinal. Then*
$\ulcorner \mathrm{V}_a(X)\urcorner$ *is* $\ulcorner\{x \mid \exists z(z \in \mathrm{M} \wedge z \subseteq X \wedge x \in \mathrm{V}_a(z))\}\urcorner$,
$\ulcorner \mathrm{V}(X)\urcorner$ *is* $\ulcorner\{x \mid \exists a(a \in \mathrm{On} \wedge x \in \mathrm{V}_a(X))\}\urcorner$.

Then (68) gives that *the sets are just the sets over the non-sets.*

(72) $\vdash_1 \ulcorner \mathrm{M} = \mathrm{V}(\mathfrak{V}\setminus \mathrm{M})\urcorner$.

Our theory **ZF+Cls** does not tell us what classes there are beyond V. The theory is compatible with any reasonable collection of real classes closed under the set formation processes. One could demand that there are no real classes beyond the hereditary sets. This would mean that the axiom

(73) $\ulcorner \mathfrak{V} = \mathrm{M}\urcorner$

holds. Then we would have $\ulcorner \mathfrak{V} = \mathrm{M} = \mathrm{V}\urcorner$, the primitive notion M would be superfluous and the formulation of the axioms could be simplified (as is usually done). With axiom (73) we would get the well-known theory **ZF** of hereditary sets. By our reformulation of set theory we intended, however, to show that it is quite compatible with the intuitive notion of set that there might be non-sets, i. e. proper classes which are elements of sets.

In such a system we would have:

(74) $\ulcorner \mathfrak{V} \neq \mathrm{M}\urcorner$

and then of course $\ulcorner \mathfrak{V} \neq \mathrm{M} \neq \mathrm{V}\urcorner$ [21].

Since we have the usual set axioms, our theory **ZF+Cls** is suitable as a "working set theory". For the working mathematician it is quite possiblet o work within **ZF+Cls** and at the same treat the question of class existence *naively*, i. e. adopt any class as real if he wants to do so, e. g. if one wants to work with a functor category, then one assumes it to be real.

This would be a semiaxiomatic treatment. Questions of set existence and sethood are decided in accordance with the axioms, questions of class existence are decided naively and ad hoc. This would correspond to the

---

[21] If we admit urelements, then $\ulcorner \mathrm{V}(\mathrm{UR})\urcorner$, Cls would play the rôle of V, $\mathfrak{V}$ in (73), (74).

fact that for sets there *is* a widely accepted system of axioms([22]) and for classes there is *no* such system of axioms. But of course, this seminaive standpoint is rather unsatisfactory. It would be better to have some fixed body of axioms of class existence in addition to **ZF+Cls**.

In the following sections we present several class existence schemes which have the form of comprehension principles([23]).

5. The M-Comprehension schema

A common opinion about classes is that any condition about sets determines a real class. This would mean that one adopts the following axiom schema of class existence:

(75) *Tentative axiom schema*: $\ulcorner X \subseteq \mathrm{M} \to X \in \mathfrak{V}\urcorner$.

Unfortunately, (75) gives a contradiction([24]) with **ZF+Cls**.

Let X be $\ulcorner\{\{x\} \mid \{x\} \notin x\}\urcorner$. Then

$$\ulcorner\{y\} \in \mathfrak{V} \vdash_1 \ulcorner\{y\} \in X \leftrightarrow \exists x(\{y\} = \{x\} \wedge \{x\} \notin x)\urcorner.$$

The premise $\ulcorner\{y\} \in \mathfrak{V}\urcorner$ can be detached because of **A 1**. Moreover, $\ulcorner\{y\} = \{x\}\urcorner$ is equivalent to $\ulcorner y = x\urcorner$, and the right-hand side of the equivalence is equivalent to $\ulcorner\{y\} \notin y\urcorner$. So we have:

$$\vdash_1 \ulcorner\{y\} \in X \leftrightarrow \{y\} \notin y\urcorner.$$

Substitution yields:

$$\ulcorner X \in \mathfrak{V}\urcorner \vdash_1 \ulcorner\{X\} \in X \leftrightarrow \{X\} \notin X\urcorner.$$

So we get: $\vdash_1 \ulcorner X \notin \mathfrak{V}\urcorner$.

But on the other hand, from **A 1** it follows $\ulcorner X \subseteq \mathrm{M}\urcorner$. So, instead of (75), we have in **ZF+Cls** the theorem:

(76) *There are formulas φ such that* $\vdash_1 \ulcorner\{x \mid x \in \mathrm{M} \wedge \varphi\} \notin \mathfrak{V}\urcorner$.

It is possible to save (75) by admitting ultimate classes (and thus virtual sets). This is in our opinion the main motivation for systems with ultimate classes. We, however, decided for **A 1** and **S 5**. Of course, we could

([22]) It is another story that the axioms of set theory are still "very" incomplete.

([23]) Perhaps it is possible to have class existence axioms directly in the form of category existence axioms. But we have the feeling that the set theoretical part is better not given in a categorical language.

([24]) This contradiction had been for some time an obstacle for us in finding the theory **ZF+M-Comp** to be presented below. The contradiction is also mentioned in the first edition of [25], p. 322, where Quine contributes the idea to Myhill.

also regain (75) in our systems by introducing ultimate classes. The old real classes would become the elements. Any old virtual class would become real and ultimate. However, we see no point in this, since the job of ultimate classes can be well done by virtual classes[25].

We saw that it is impossible that any condition on sets defines a real class. *But it is possible that any set theoretical condition defines a real class.* A set theoretical condition is a formula speaking only about sets, i. e. a simple formula having only bound set variables[26].

(77) *It is relative consistent to adjoin to* **ZF+Cls** *all axioms:*

$$\ulcorner x_1, \ldots, x_n \in \mathrm{M} \to \{x \mid x \in \mathrm{M} \wedge \varphi\} \in \mho\urcorner,$$

where $x_1, \ldots, x_n, x$ *are the free variables of* φ *and* φ *is a set theoretical condition, i. e. a simple formula with all bound variables in* φ *relativized to* M.

Actually, a much more general comprehension principle can be adjoined to **ZF+Cls** without giving contradictions (if **ZF** is consistent). This has been shown in the paper [14][27].

For convenience we give here again in a different terminology what is essentially the system $\mathbf{G}^*$ of [14]. The comprehension principle will be called M-*comprehension schema* and the resulting system will be called **ZF+M-Comp.**

The key notion in the M-comprehension schema is a certain relativization process. A variable *x is relativized to* M in a formula φ if all parts $\ulcorner\forall x \ldots\urcorner$, $\ulcorner\exists x \ldots\urcorner$, $\ulcorner\{x \mid \ldots\}\urcorner$, $\ulcorner\iota x \ldots\urcorner$ are replaced by $\ulcorner\forall x(x \in \mathrm{M} \to \ldots)\urcorner$, $\ulcorner\exists x(x \in \mathrm{M} \wedge \ldots)\urcorner$, $\ulcorner\{x \mid x \in \mathrm{M} \wedge \ldots\}\urcorner$, $\ulcorner\iota x(x \in \mathrm{M} \wedge \ldots)\urcorner$.

We call a formula or general term *usual* if bound variables with different scopes are different and if all bound variables are different from all free variables. By alphabetic variance any formula or term is logically equivalent to a usual formula or term.

---

(25) If one really wants to add ultimate classes, then we would prefer to use different styles of variables (allowing equalities between both sorts of variables) for elements and for objects and call the resulting system the *second order version* of the original system. Thus $\mathbf{ZF}^{(2)}$ the second order **ZF**-set theory is the system introduced for the first time by Wang [31] and known under several names (**NQ**, Morse's set theory, Kelley's set theory, **NBGQ**, **BQ**), and $\mathbf{NF}^{(2)}$ is the same as Quine's system **ML**. The second order versions of our systems could easily be presented.

(26) The condition that φ is simple can not be dropped completely. Note that in our flexible language $\ulcorner X \in Y\urcorner$ may show only bound set variables. But actually there is hidden another variable which becomes apparent in the reformulation $\ulcorner\exists x(x = X \wedge x \in Y)\urcorner$. So $\ulcorner X \in Y\urcorner$ is not intuitively a set theoretical condition (rather $\ulcorner X \in \mathrm{M} \wedge X \in Y\urcorner$ is).

(27) Actually we proved this only for a theory with Aussonderungsaxiom and Ersetzungsaxiom instead of the schemas. But R. B. Jensen pointed out that the schemas can also be shown relative consistent. See section 6.

We recall the definition of simple formulas (see (13)).

(78) *Let φ be a usual and simple formula. Then φ^{M} is any formula that comes from φ by relativizing certain variables to* M. *The variables which are to be relativized to* M *are determined in the following way. To all variable occurrences in φ (other than immediately after a quantifier) is given an index which is either* 0 *or* 1 *in such a way that:*

(i) *any occurrence immediately to the left of ϵ gets index* 0 *and any occurrence immediately to the right of ϵ gets index* 1,

(ii) *the indicated occurrences of x and y in a subformula $\ulcorner x = y \urcorner$ get the same index,*

(iii) *the indicated occurrences of z, x, and y in a subformula $\ulcorner z = \langle x, y\rangle \urcorner$ get index* 0,

(iv) *The indicated occurrences of z and x in a subformula $\ulcorner z = \iota x \psi \urcorner$ get the same index,*

(v) *the indicated occurrences of z and x in a subformula $\ulcorner z \epsilon \{x \mid \psi\} \urcorner$ get index* 0,

(vi) *the indicated occurrence of z in a subformula $\ulcorner z = \{x \mid \psi\} \urcorner$ gets index* 1 *and the indicated occurrence of x gets index* 0.

Then all variables getting both indices 0 *and* 1 *(at different occurrences of course) are to be relativized to* M *unless they are already relativized to* M. *The resulting formula φ^{M} is called a* M*-formula.*

We give some examples of M-formulas. An indexing according to (78) is given by superscripts to the variables.

$$\ulcorner x^0 = y^0 \urcorner,\ \ulcorner x^1 = y^1 \urcorner,\ \ulcorner z^0 = \langle x^0, y^0\rangle \urcorner,\ \ulcorner \forall u (u^0 \epsilon x^1 \to u^0 \epsilon y^1) \urcorner,$$

$$\ulcorner \exists w (w^0 = \langle u^0, v^0\rangle \wedge w^0 \epsilon x^1) \urcorner,\ \ulcorner \forall w (w^0 \epsilon x^1 \to \exists u \exists v\, w^0 = \langle u^0, v^0\rangle) \urcorner,$$

$$\ulcorner \forall w (w^0 \epsilon x^1 \to \exists u \exists v\, w^0 = \langle u^0, v^0\rangle) \wedge \forall u_1 \forall u_2 \forall v (\exists w_1 \exists w_2 (w_1^0 = \langle u_1^0, v^0\rangle$$

$$\wedge\, w_2^0 = \langle u_2^0, v^0\rangle \wedge w_1^0 \epsilon x^1 \wedge w_2^0 \epsilon x^1) \to u_1^0 = u_2^0) \urcorner,$$

$$\ulcorner z^0 = \iota u^0 \exists w (w^0 = \langle u^0, y^0\rangle \wedge w^0 \epsilon x^1) \urcorner.$$

The M-*comprehension schema* is:

(79) **M-Comp**: *Let φ be usual and simple and z be new. Then the following is an axiom* ([28]):

$$\ulcorner \forall \ldots \exists z\, z = \{u \mid \varphi\} \urcorner^{\mathrm{M}}.$$

([28]) $\ulcorner \forall \ldots \theta \urcorner$ is the universal closure of θ.

ZF+M-Comp is the theory which has the axioms of **ZF+Cls** and in addition all axioms of **M-Comp**. We write $\Sigma \vdash_2 \varphi$ to indicate $\Sigma \cup A \vdash \varphi$ for some set A of (closed) formulas which are axioms of **ZF+M-Comp**. Thus $\vdash_2 \varphi$ means that φ is a theorem of **ZF+M-Comp**. This system is relative consistent to a version of **ZF** with urelements (see [13]), and also to **ZF**.

The following is an example of an axiom:

$$\ulcorner \exists z\, z^1 = \{u^0 \mid u^0 \in \mathrm{M}\}\urcorner.$$

We have given indices as superscripts in order to show that this is a M-formula. Since $\vdash_1 \ulcorner \mathrm{M} = \{u \mid u \in \mathrm{M}\}\urcorner$, we have the theorem:

(80) $\quad \vdash_2 \ulcorner \mathrm{M} \in \mathfrak{V}\urcorner.$

This can not be proved in **ZF+Cls**, and therefore **ZF+M-Comp** is a proper extension. Since M is a proper class, we also have:

(81) $\quad \vdash_2 \ulcorner \mathrm{PC} \neq \emptyset\urcorner.$

We list some more instances of the M-comprehension schema. We always give indices in order to show that the formula is a M-formula:

$$\ulcorner \exists z\, z^1 = \{u^0 \mid u^0 = u^0\}\urcorner, \qquad \ulcorner \exists z\, z^1 = \{u^0 \mid u^0 \neq u^0\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{u^0 \mid u^0 = x^0 \vee u^0 = y^0\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{u^0 \mid u^0 \in x^1 \wedge u^0 \in y^1\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{u^0 \mid u^0 \in x^1 \vee u^0 \in y^1\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{u^0 \mid u^0 \notin x^1\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{u^0 \mid \exists y\,(y^0 \in \mathrm{M} \wedge u^0 \in y^1 \wedge y^0 \in x^1)\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{u^0 \mid \forall y\,(y^0 \in \mathrm{M} \to (y^0 \in x^1 \to u^0 \in y^1))\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{u^0 \mid u^0 \in \mathrm{M} \wedge \forall y\,(y^0 \in u^1 \to y^0 \in x^1)\}\urcorner.$$

Therefore we have the theorem:

(82) $\quad \vdash_2 \ulcorner \mathfrak{V}, \emptyset, \{x, y\}, x \cap y, x \cup y, -x, \bigcup x, \bigcap x, \mathrm{P}x \in \mathfrak{V}\urcorner.$

This shows that the axioms **A 0, A 1, A 2, A 3**, in **ZF+M-Comp** are rather set*hood*-axioms (of classes which exist according to the comprehension schema) than set-*existence*-axioms.

The universal class is also closed under the operations of the elementary theory of relations. In order to see this, we observe that the formula $\ulcorner uxv\urcorner$ (which is the same as $\ulcorner\langle u, v\rangle \in x\urcorner$ and which was already considered under

(78)) is a M-formula with variables u, v getting index 0, and x getting index 1. Therefore the following formulas are axioms:

$$\ulcorner \exists z\, z^1 = \{w^0 \mid \exists u\, \exists v (w^0 = \langle u^0, v^0\rangle \wedge u^0 \in x^1 \wedge v^0 \in y^1)\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{w^0 \mid \exists u\, \exists v (w^0 = \langle u^0, v^0\rangle \wedge u^0 = v^0 \wedge v^0 \in x^1)\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{w^0 \mid \exists u\, \exists v (w^0 = \langle u^0, v^0\rangle \wedge v^0 x^1 u^0)\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{w^0 \mid \exists u\, \exists v \big(w^0 = \langle u^0, v^0\rangle \wedge \exists s (u^0 x^1 s^0 \wedge s^0 y^1 v^0)\big)\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{w^0 \mid \exists u\, \exists v (w^0 = \langle u^0, v^0\rangle \wedge u^0 x^1 v^0 \wedge v^0 \in y^1)\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{v^0 \mid \exists u\, u^0 x^1 v^0\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \{u^0 \mid \exists v\, u^0 x^1 v^0\}\urcorner.$$

This gives the theorem

(83) $\quad \vdash_2 \ulcorner x \times y, \mathrm{id}_x, x^{-1}, x \circ y, x \restriction y, \mathrm{dom}\, x, \mathrm{ran}\, x \in \mathfrak{V}\urcorner.$

Next we show that the class of hereditary sets is real. It holds

$$\vdash_1 \ulcorner \exists z\, z = \mathrm{V} \leftrightarrow \exists z\, z = \{x \mid x \in \mathrm{M} \wedge \mathrm{TC}(x) \subseteq \mathrm{M}\}\urcorner.$$

The right-hand side will be transformed equivalently:

$$\ulcorner \exists z\, z = \{x \mid x \in \mathrm{M} \wedge \forall y \big(y \in \mathrm{TC}(x) \to y \in \mathrm{M}\big)\}\urcorner,$$
$$\ulcorner \exists z\, z = \Big\{x \mid x \in \mathrm{M} \wedge \forall y \Big(\forall z \big(z \in \mathrm{M} \to (\mathrm{Trans}_{\mathrm{M}} z \wedge x \subseteq z \to y \in z)\big) \to y \in \mathrm{M}\Big)\Big\}\urcorner,$$
$$\ulcorner \exists z\, z^1 = \Big\{x^0 \mid x^0 \in \mathrm{M} \wedge \forall y \Big(\forall z \Big(z^0 \in \mathrm{M} \to \Big(\forall u \forall v \big(v^0 \in \mathrm{M} \to (u^0 \in v^1 \wedge v^0 \in z^1$$
$$\to u^0 \in z^1)\big) \wedge \forall w (w^0 \in x^1 \to w^0 \in z^1) \to y^0 \in z^1\Big)\Big) \to y^0 \in \mathrm{M}\Big)\Big\}\urcorner.$$

We have given superscripts in the last formula. It can easily be verified that variables x, z, v are correctly relativized to M and that the other variables need no relativization. Therefore the last formula is axiom and we have the theorem:

(84) $\quad \vdash_2 \ulcorner \mathrm{V} \in \mathfrak{V}\urcorner.$

We called a *set theoretical condition* a simple and usual formula such that all variables are relativized to M. This is certainly a M-formula and therefore we have the theorem (77) that these conditions define real classes.

Let φ be a *condition purely on* V with free variables $x, x_1, \ldots, x_n$ (i. e. φ is simple and usual and all bound variables are relativized to V). This is equivalent to a M-formula since the relativization clauses $\ulcorner v \in \mathrm{V}\urcorner$ are equivalent to M-formulas (giving index 0 to v) and the whole formula can undergo equivalently an additional relativization to M. Thus φ will define a real class.

(85) $\ulcorner x_1, \ldots, x_n \in \mathrm{V}\urcorner \vdash_2 \ulcorner\{x \mid x \in \mathrm{V} \wedge \varphi\} \in \mathfrak{V}\urcorner$ (*if* φ *is purely on* V *with free* $x_1, \ldots, x_n x$).

Therefore we can say that all classes of the wellknown system **NBG** exist in **ZF+M-Comp.**

Actually, for the last theorem it would be convenient to identify the primitive ordered pair $\langle x, y\rangle$ with $\{\{x\}, \{x, y\}\}$ which maps sets onto sets and hereditary sets onto hereditary sets. But this is possible since the formula $\ulcorner z = \{\{x\}, \{x, y\}\}\urcorner$ is equivalent to a M-formula with z, x, y getting index 0. Since this is the only property of $\ulcorner z = \langle x, y\rangle\urcorner$ we need here, we can identify $\langle x, y\rangle$ and $\{\{x\}, \{x, y\}\}$. We give an equivalent reformulation of $\ulcorner z = \{\{x\}, \{x, y\}\}\urcorner$ with indices now:

$$\ulcorner \exists z_1 \Big(z_1^0 \in \mathrm{M} \wedge z_1^0 = z^0 \wedge \forall w_1 (w_1^0 \in z_1^1 \to w_1^0 \in \mathrm{M}) \vee \forall w_2 \Big(w_2^0 \in \mathrm{M} \to \big(w_2^0 \in z_1^1$$
$$\leftrightarrow \big(\forall u (u^0 \in w_2^1 \leftrightarrow u^0 = x^0) \vee \forall v (v^0 \in w_2^1 \leftrightarrow v^0 = x^0 \vee v^0 = y^0)\big)\big)\Big)\Big)\urcorner.$$

The theory **ZF+M-Comp** thus proves to be quite comprehensive. One could ask whether in this theory there are "big structures", e. g. as the permutation group of $\mathfrak{V}$. But it seems not possible to derive in **ZF+M-Comp** that the class of all (real) one-one-function from $\mathfrak{V}$ to $\mathfrak{V}$ is real.

In section 7 we present the system **ZF+NF** where classes like this exist.

6. Further considerations on ZF+M-Comp

Before going to the next set theory over classes, i. e. to **ZF+NF**, we bring some additional material about **ZF+M-Comp.**

It is tiresome to have the M-comprehension schema only for usual and simple formulas. For practical work it would be convenient to work with the full language and to use abbreviations.

Our first topic will be to accomplish this. We introduce certain well-behaved formulas and terms to which we assign signatures.

We will give rules how these formulas and terms can be combined and how a signature of the result is computed. In doing so, one can work with abbreviations. Any well-behaved term will denote a real class.

Signatures are defined by means of three symbols which we write: $0, 1, \mathrm{M}$.

(86) *A signature σ is a mapping of some finite set of variables into $\{0, 1, \mathrm{M}\}$. The variables in the domain of σ are the variables of σ, the variables x of σ with $\sigma(x) = 0$ are the* E*-variables, with $\sigma(x) = 1$ are the* C*-variables, and with $\sigma(x) = \mathrm{M}$ are the* M*-variables* of σ. *If the domain of σ is $\{x_1, \ldots, x_n\}$ (with $x_1, \ldots, x_n$ different) and $\sigma(x_k) = i_k$ $(1 \leqslant k \leqslant n)$, then we say that σ is $\langle i_1, \ldots, i_n\rangle$ with respect to $x_1, \ldots, x_n$.*

E. g. if σ is $\langle 0, 0, \mathrm{M}\rangle$ with respect to x, y, z, then x, y are the E-variables, z is the (only) M-variable, there are no C-variables.

(87) *Let σ be a signature and the free variables of φ be variables of σ. φ is a* M*-expression of signature σ iff there is a simple and usual* M*-formula ψ containing exactly the variables of σ free such that under some assignment of indices in ψ according to* (78) *for any x of σ:*

(i) $\sigma(x) = 0$ *iff x is getting index* 0 *throughout ψ,*

(ii) $\sigma(x) = 1$ *iff x is getting index* 1 *throughout ψ,*

(iii) $\sigma(x) = \mathrm{M}$ *iff x is getting index* 0 *at some places in ψ and index* 1 *at other places in ψ,*

(iv) *if $z_1, \ldots, z_l$ are the* M*-variables of σ, then $\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner \varphi \leftrightarrow \psi\urcorner$.*

If σ is $\langle i_1, \ldots, i_n\rangle$ with respect to $x_1, \ldots, x_n$, then we also say that φ is a M-expression of signature $\langle i_1, \ldots, i_n\rangle$ with respect to $x_1, \ldots, x_n$. The reference to the variables is omitted if it is clear from the context.

We also speak of the E-variables, C-variables, M-variables of a formula or term. This is, of course, always meant with respect to some specified signature. M-expressions are thus formulas equivalent (in **ZF+M-Comp**) to usual and simple M-formulas. The signatures show which indices are given to the free variables. A M in the signature indicates that both indices are given to the corresponding variable. The equivalence need only hold if these M-variables are restricted to sets.

The following formulas are M-expressions of the given signatures. It is always understood (here and in future) that the signature is with respect to the free variables in the order of first occurrence in the abbreviation, given for the formula:

$\ulcorner x \in y\urcorner$ *Signature*: $\langle 0, 1\rangle$,
$\ulcorner x = y\urcorner$ *Signatures*: $\langle 0, 0\rangle, \langle 1, 1\rangle$,
$\ulcorner z = \langle x, y\rangle\urcorner$ *Signature*: $\langle 0, 0, 0\rangle$,
$\ulcorner x \subseteq y\urcorner$ *Signature*: $\langle 1, 1\rangle$,
$\ulcorner uxv\urcorner$ *Signature*: $\langle 0, 1, 0\rangle$,
$\ulcorner \mathrm{Relation}(x)\urcorner$, $\ulcorner \mathrm{Function}(x)\urcorner$ *Signature*: $\langle 1\rangle$.

The equivalent M-formulas have already been given.

Signatures are not uniquely determined. To any formula φ we can equivalently add by conjunction clauses like $\ulcorner\exists y(x \in y \vee x \notin y)\urcorner$ or $\ulcorner\exists y(y \in x \vee y \notin x)\urcorner$ or $\ulcorner x \in x \vee x \notin x\urcorner$ bringing in additional occurrences of x either with index 0, or index 1, or both indices. This shows that we can change in a signature a 0 or 1 into a M and that we can extend signatures with resprect to additional variables in an arbitrary way. But in general it is not possible to change a M into 0 or 1 or to interchange 0 and 1.

The *superposition* of the signatures σ, τ is the function ϱ coinciding with σ on the domain of σ and coinciding with ϱ on the domain of ϱ, except for the case that σ and τ differ for a variable x which is the domain of σ and in the domain of τ. In that case $\varrho(x) = \mathrm{M}$.

E. g. if σ is $\langle 1, 0, 1\rangle$ with respect to z, x, y and τ is $\langle 1, \mathrm{M}, 0\rangle$ with respect to y, u, z, then the superposition is $\langle 0, 1, \mathrm{M}, \mathrm{M}\rangle$ with respect to x, y, z, u.

Deleting a variable x of a signature σ means to restrict σ to the domain from which x is deleted.

E. g. if σ is $\langle 0, 1, \mathrm{M}, \mathrm{M}\rangle$ with respect to x, y, z, u, then $\langle 0, 1, \mathrm{M}\rangle$ with respect to x, y, u is got by deleting z.

The composition of M-expression is governed by the following theorem:

(88) *If* φ, ψ *are* M-*expressions, then* $\ulcorner\neg\varphi\urcorner$, $\ulcorner(\varphi\wedge\psi)\urcorner$, $\ulcorner(\varphi\vee\psi)\urcorner$ *are* M-*expressions.*
If φ *is a* M-*expression and* x *is a* E-*variable or* C-*variable, then* $\ulcorner\forall x\varphi\urcorner$, $\ulcorner\exists x\varphi\urcorner$ *are* M-*expressions. If* φ *is a* M-*expression and* x *is arbitrary, then* $\ulcorner\forall x(x \in \mathrm{M} \to \varphi)\urcorner$, $\ulcorner\exists x(x \in \mathrm{M} \wedge \varphi)\urcorner$ *are* M-*expressions. In any case it is clear, what is the resulting signature.*

For the proof of (88) suppose

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\varphi \leftrightarrow \varphi'\urcorner, \quad \ulcorner y_1, \ldots, y_k \in \mathrm{M}\urcorner \vdash_2 \ulcorner\psi \leftrightarrow \psi'\urcorner$$

for M-formulas φ', ψ' which can be considered to be without common bound variables. Then we have

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\neg\varphi \leftrightarrow \neg\varphi'\urcorner$$

$$\ulcorner z_1, \ldots, z_l, y_1, \ldots, y_k \in \mathrm{M}\urcorner \vdash_2 \ulcorner\big((\varphi\wedge\psi) \leftrightarrow (\varphi'\wedge\psi')\big) \wedge \big((\varphi\vee\psi) \leftrightarrow (\varphi'\vee\psi')\big)\urcorner.$$

The right-hand sides of the equivalences are M-formulas.

If x is not a M-variable, i. e. not among $z_1, \ldots, z_l$, then we can derive:

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner(\forall x\varphi \leftrightarrow \forall x\varphi')\urcorner \wedge (\exists x\varphi \leftrightarrow \exists x\varphi')\urcorner.$$

The right-hand sides of the equivalences are M-formulas (give x the same index as in φ').

If x is a M-variable, say x is z_1, then we can derive

$$\ulcorner z_2, \ldots, z_l \in M\urcorner \vdash_2 \ulcorner\big(\forall x(x \in \mathrm{M} \to \varphi) \leftrightarrow \forall x(x \in \mathrm{M} \to \varphi')\big)$$
$$\wedge\big(\exists x(x \in \mathrm{M} \wedge \varphi) \leftrightarrow \exists x(x \in \mathrm{M} \wedge \varphi')\big)\urcorner.$$

The right-hand sides of the equivalences are M-formulas (give to the indicated occurrence of x index 0).

If x is not a M-variable, we can make it a M-variable by changing the signature and proceed as before.

The signature of $\ulcorner \neg \varphi\urcorner$ is the same as of φ.

The signature of $\ulcorner(\varphi \wedge \psi)\urcorner$, $\ulcorner(\varphi \vee \psi)\urcorner$ is the superposition of the signatures of φ and of ψ.

Signatures of the quantified formulas are got by deleting x in a signature of φ.

We now introduce E-*terms* and C-*terms*.

(89) *Let τ, σ be signatures, z a variable of τ, $\tau(z) = 0$ or $\tau(z) = 1$, let σ be got from τ by deleting z, let $z_1, \ldots, z_l$ be the* M-*variables of σ, and the free variables of X be variables of σ. Suppose that*

$$\ulcorner z = X\urcorner$$

is a M-*expression of signature τ and:*
$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner X \in \mathcal{V}\urcorner$. *Then:*
X is a E-*term of signature σ iff $\tau(z) = 0$,*
X is a C-*term of signature σ iff $\tau(z) = 1$.*

Let us use "*regular*" as a common denotation for M-expressions, E-terms, and C-terms.

It is convenient to write the signature of a regular term rather as a $(n+1)$-tuple than as a n-tuple, adding a first component, indicating whether it is a E-term or C-term. Therefore, if σ is $\langle i_1, \ldots, i_n\rangle$ with respect to $x_1, \ldots, x_n$, then we say that X is of signature $\langle 0 \,|\, i_1, \ldots, i_n\rangle$ with respect to $x_1, \ldots, x_n$ iff X is a E-term of signature σ, and X is of signature $\langle 1 \,|\, i_1, \ldots, i_n\rangle$ with respect to $x_1, \ldots, x_n$ iff X a C-term of signature σ.

We give examples of regular terms. The equivalents of formulas $\ulcorner z = X\urcorner$ with a suitable indexing have already occurred, and $\ulcorner X \in \mathcal{V}\urcorner$ is already known.

x, *signatures*: $\langle 0\,|0\rangle$, $\langle 1\,|1\rangle$,
$\ulcorner\langle x, y\rangle\urcorner$, *signature*: $\langle 0\,|0, 0\rangle$,
$\mathrm{M}, \emptyset, \mathcal{V}$, *signature*: $\langle 1\,|\rangle$,
$\ulcorner\{x, y\}\urcorner$, *signature*: $\langle 1\,|0, 0\rangle$,
$\ulcorner x \cap y\urcorner$, $\ulcorner x \cup y\urcorner$, $\ulcorner x \times y\urcorner$, $\ulcorner x \circ y\urcorner$, $\ulcorner x \restriction y\urcorner$, *signature*: $\langle 1\,|1, 1\rangle$,
$\ulcorner -x\urcorner$, $\ulcorner\bigcup x\urcorner$, $\ulcorner\bigcap x\urcorner$, $\ulcorner \mathrm{P}x\urcorner$, $\ulcorner \mathrm{id}_x\urcorner$, $\ulcorner x^{-1}\urcorner$, $\ulcorner \mathrm{dom}\, x\urcorner$, $\ulcorner \mathrm{ran}\, x\urcorner$, *signature*: $\langle 1\,|1, 1\rangle$,
$\ulcorner x(y)\urcorner$, *signature*: $\langle 0\,|1, 0\rangle$.

Regular terms that denote sets are very convenient since they can serve both as C-terms and E-terms.

(90) *Let X be a regular term with the* M-*variables* $z_1, \ldots, z_l$.
Suppose $\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner X \in \mathrm{M}\urcorner$.
Then X is as well a C-*term as a* E-*term.*

Then we can derive:

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner z = X \leftrightarrow \exists y (y \in \mathrm{M} \wedge y = z \wedge y = X)\urcorner.$$

If X is a C-term, then we give indices to the new variables of the right-hand side of the equivalence in the following way:

$$\ulcorner \exists y (y^0 \in \mathrm{M} \wedge y^0 = z^0 \wedge y^1 = X)\urcorner.$$

This shows that X is a E-term.

If X is a E-term, then we give indices in the following way:

$$\ulcorner \exists y (y^0 \in \mathrm{M} \wedge y^1 = z^1 \wedge y^0 = X)\urcorner,$$

showing that X is a C-term.

Using this theorem and axiom **A 1**, we get a new signature for $\{x, y\}$:

$$\ulcorner\{x, y\}\urcorner \textit{ has also signature } \langle 0 | 0, 0\rangle.$$

Using **A 0**, we get a new signature of $\ulcorner\emptyset\urcorner$:

$$\ulcorner\emptyset\urcorner \textit{ has also signature } \langle 0 | \rangle.$$

The composition of regular terms is governed by (91), (92), (93).

(91) *If X, Y are* E-*terms, then* $\ulcorner\langle X, Y\rangle\urcorner$ *is a* E-*term.*
For a signature of this term take the superposition of signatures of X and Y.

For the proof consult (15) and replace the formulas $x = X$, $y = X$ occurring there by M-formulas without common bound variables. Note that $\ulcorner z = \perp\urcorner$ is now $\ulcorner z = \emptyset\urcorner$ and can be considered as a M-expression of signature $\langle 0\rangle$. $\ulcorner\langle X, Y\rangle \in \mho\urcorner$ is already stated in (12).

(92) *Suppose φ is a* M-*expression.*
If x is a E-*variable, then* $\ulcorner \iota x \varphi\urcorner$ *is a* E-*term;*
if x is a C-*variable, then* $\ulcorner \iota x \varphi\urcorner$ *is a* C-*term;*
if x is arbitrary, then $\ulcorner \iota x (x \in \mathrm{M} \wedge \varphi)\urcorner$ *is as well a* C-*term as a* E-*term.*
Signatures are got by deleting x.

Suppose $\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner \varphi \leftrightarrow \psi\urcorner$ for a M-formula ψ.

If x is a E-variable or C-variable, i. e. not among $z_1, \ldots, z_l$, then:

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner z = \iota x \varphi \leftrightarrow z = \iota x \psi\urcorner.$$

The right-hand side of the equivalence is a M-formula (give to z and the indicated occurrence of x the same index as x has in ψ).

If x is a M-variable, say x is z_1, then (using $\ulcorner\perp = \emptyset \in \mathrm{M}\urcorner$), we can derive:

$$\ulcorner z_2, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner z = \iota x(x \in \mathrm{M} \wedge \varphi) \leftrightarrow z = \iota x(x \in \mathrm{M} \wedge \psi)\urcorner.$$

Again the right-hand side of the equivalence is a M-formula (give to z and the first indicated occurrence of x the same index, which may be different from the index of x in ψ).

If x is not a M-variable, then we can make it a M-variable by changing the signature and proceed as before.

Finally we remark that $\ulcorner \iota x\varphi, \iota x(x \in \mathrm{M} \wedge \varphi) \in \mathfrak{V}\urcorner$ is already stated in (12).

Axiom **OP′** would imply that all regular pair terms are C-terms as well.

(93) *Suppose φ is a* M-*expression.*
If x is a E-*variable, then $\ulcorner\{x|\varphi\}\urcorner$ is a* C-*term.*
If x is arbitrary, then $\ulcorner\{x|x \in \mathrm{M} \wedge \varphi\}\urcorner$ is a C-*term.*
Signatures are got by deleting x.

Suppose $\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\varphi \leftrightarrow \psi\urcorner$ for a M-formula ψ.

If x is a E-variable and thus not among $z_1, \ldots, z_l$, then we can derive:

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner z = \{x|\varphi\} \leftrightarrow z = \{x|\psi\}\urcorner.$$

The right-hand side of the equivalence is a M-formula (give index 1 to z and index 0 to x).

If x is a M-variable, say x is z_1, then we can derive:

$$\ulcorner z_2, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner z = \{x|x \in \mathrm{M} \wedge \varphi\} \leftrightarrow z = \{x|x \in \mathrm{M} \wedge \psi\}\urcorner.$$

Again the right-hand side of the equivalence is a M-formula (give index 1 to z and index 0 to the indicated occurrences of x).

If x is not a M-variable, then we can make it a M-variable by changing the signature and proceed as before.

The M-comprehension schema yields:

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\{x|\psi\} \in \mathfrak{V}\urcorner \quad \text{(in the first case)},$$

$$\ulcorner z_2, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\{x|x \in \mathrm{M} \wedge \psi\} \in \mathfrak{V}\urcorner \quad \text{(in the second case)}.$$

Therefore we also have:

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\{x|\varphi\} \in \mathfrak{V}\urcorner \quad \text{(in the first case)},$$

$$\ulcorner z_2, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\{x|x \in \mathrm{M} \wedge \varphi\} \in \mathfrak{V}\urcorner \quad \text{(in the second case)}.$$

Now we give substitution properties for regular expressions and terms:

(94) *Suppose φ is a* M-*expression and X a regular term.*

(i) *If x is a* E-*variable of φ and X is a* E-*term, or if x is a* C-*variable of φ and X is a* C-*term, then φ^x_X is a* M-*expression.*

(ii) *If x and X are arbitrary, then*

$$\ulcorner X \in \mathrm{M} \wedge \varphi^x_X \urcorner$$

is a M-*expression.*

A signature is got if we delete x in the signature of φ and superpose it with the signature of X.

Let $z_1, \ldots, z_l$ be the M-variables of X. Then we have $\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner X \in \mho\urcorner$. Therefore:

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner \varphi^x_X \leftrightarrow (\exists z z = X \wedge \varphi^x_z)\urcorner.$$

The right-hand side of the equivalence is a M-expression. This is seen when we give to z in $\ulcorner z = X\urcorner$ and φ^x_z the same index as x has in φ.

We also can derive:

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner X \in \mathrm{M} \wedge \varphi^x_X \leftrightarrow \exists z (z \in \mathrm{M} \wedge z = X \wedge \varphi^x_z)\urcorner.$$

Again the right-hand side of the equivalence is a M-expression. This is seen when we give to z in $\ulcorner z \in \mathrm{M}\urcorner$ index 0, to z in $\ulcorner z = X\urcorner$ index 0 or 1 (as X is a E-term or C-term) and to z in φ^x_z everywhere the same index as x has at corresponding places in φ.

We can also substitute into terms:

(95) *Suppose Y, X are regular terms.*

(i) *If x is a* E-*variable of Y and X is a* E-*term, or if x is a* C-*variable of Y and X is a* C-*term, then Y^x_X is a regular term.*

(ii) *If x is arbitrary and $\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner X \in \mathrm{M}\urcorner$, where $z_1, \ldots, z_l$ are the* M-*variables of x, then Y^x_X is a regular term.*

In any case Y^x_X is a E-*term or* C-*term according to whether Y is a* E-*term or* C-*term. A signature is got if we delete x in the signature of Y and superpose it with the signature of X.*

The proof that $z = Y^x_X$ is a M-expression is simple by applying (94) to the formula $z = Y$. In case (ii) use the additional hypothesis to derive:

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner z = Y^x_X \leftrightarrow X \in \mathrm{M} \wedge z = Y^x_X\urcorner.$$

Finally we have to prove the reality of Y^x_X. Let $y_1, \ldots, y_m$ be the M-variables of Y. By regularity of X and y we have:

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner X \in \mho\urcorner \text{ and } \ulcorner y_1, \ldots, y_m \in \mathrm{M}\urcorner \vdash_2 \ulcorner Y \in \mho\urcorner.$$

Since in case (i) x is not among $y_1, \ldots, y_m$, we even have:

$$\ulcorner y_1, \ldots, y_m \in \mathrm{M}\urcorner \vdash_2 \ulcorner \forall x Y \in \mathfrak{V}\urcorner.$$

Since $\ulcorner X \in \mathfrak{V}\urcorner, \ulcorner \forall x Y \in \mathfrak{V}\urcorner \vdash \ulcorner Y^x_X \in \mathfrak{V}\urcorner$, we get:

$$\ulcorner z_1, \ldots, z_l, y_1, \ldots, y_m \in \mathrm{M}\urcorner \vdash_2 \ulcorner Y^x_X \in \mathfrak{V}\urcorner.$$

The M-variables of Y^x_X certainly comprise $z_1, \ldots, z_l, y_1, \ldots, y_m$.

In case (ii) we may assume that x is y_1. Then we get

$$\ulcorner y_2, \ldots, y_m \in \mathrm{M}\urcorner \vdash_2 \ulcorner \forall x(x \in \mathrm{M} \to Y \in \mathfrak{V})\urcorner.$$

Since

$$\ulcorner X \in \mathrm{M}\urcorner, \ulcorner \forall x(x \in \mathrm{M} \to Y \in \mathfrak{V})\urcorner \vdash \ulcorner Y^x_X \in \mathfrak{V}\urcorner,$$

we get

$$\ulcorner z_1, \ldots, z_l, y_2, \ldots, y_m \in \mathrm{M}\urcorner \vdash_2 \ulcorner Y^x_X \in \mathfrak{V}\urcorner.$$

The M-variables of Y^x_X certainly comprise $z_1, \ldots, z_l, y_2, \ldots, y_m$.

Theorems (88) (*composition of regular expressions*), (91) (*composition of regular pair terms*), (92) (*composition of regular description terms*), (93) (*composition of regular abstraction terms*), (94) (*substitution into regular expressions*), and (95, i) (*substitution into regular terms*) give *purely syntactical* conditions for composing regular expressions and terms and computing their signatures. The important point is that one need not expand formulas into primitive notation and transform them into simple formulas. One can work with abbreviations and need only keep record of the signatures which is quite easy. Of course, we are interested in regular expressions and terms because of the following theorem (which in our presentation is nothing but a part of the definition of regular terms):

(96) *If X is any regular abstraction term with* M-*variables* $z_1, \ldots, z_l$, *then*

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner X \in \mathfrak{V}\urcorner.$$

In theorem (96) *together with the syntactical theorems about the composition of regular expressions and terms, one can see a sort of reformulation of the* M-*comprehension principle which is more convenient for practical work.*

An additional flexivity is given by (90) and (95, ii). If we have proved that a regular term X denotes a set, then it can be used in a rather arbitrary way for the composition of other regular expressions and terms. (Also in (94, ii) then clause $\ulcorner X \in \mathrm{M}\urcorner$ can be deleted.) We bring some examples for the substitution.

$\ulcorner x \subseteq y \urcorner$ is of signature $\langle 1, 1\rangle$, $\ulcorner \mathrm{ran}\, z \urcorner$ is of signature $\langle 1|1\rangle$. So we can substitute $\ulcorner \mathrm{ran}\, z \urcorner$ for x in $\ulcorner x \subseteq y \urcorner$ and get $\ulcorner \mathrm{ran}\, z \subseteq y \urcorner$ of signature $\langle 1, 1\rangle$ (with respect to z, y).

Combining it with $\ulcorner \mathrm{function}(z) \urcorner$ (of signature $\langle 1\rangle$) and $\ulcorner \mathrm{dom}\, z = x \urcorner$ (of signature $\langle 1, 1\rangle$) to the formula $\ulcorner \mathrm{function}(z) \wedge \mathrm{dom}\, z = x \wedge \mathrm{ran}\, z \subseteq y \urcorner$, we see:

$\ulcorner z | x \to y \urcorner$ is a M-expression with signature $\langle 1, 1, 1\rangle$.

Substituting $\ulcorner \{x\} \urcorner$ (of signature $\langle 0|0\rangle$ for u and $\ulcorner \{x, y\} \urcorner$ (of signature $\langle 0|0, 0\rangle$) for v into $\ulcorner \{u, v\} \urcorner$ (of signature $\langle 0|0, 0\rangle$) we see:

$\ulcorner \{\{x\}, \{x, y\}\} \urcorner$ is a E-term of signature $\langle 0|0, 0\rangle$.

A corresponding M-formula had already been given at the end of section 5.

We now give some more general applications.

Consider the formula $\ulcorner z = X \wedge \varphi \urcorner$ for regular X, φ and new z.

This will have certain M-variables and $y_1, \ldots, y_m$ may be the M-variables among $x_1, \ldots, x_n$ and $z_1, \ldots, z_l$ may be the remaining M-variables. Then

$$\ulcorner \exists x_1 \ldots \exists x_n (y_1, \ldots, y_m \in \mathrm{M} \wedge z = X \wedge \varphi) \urcorner$$

is regular with M-variables $z_1, \ldots, z_l$.

If X is a E-term, then z is a E-variable and

$$\ulcorner \{z | \exists x_1 \ldots \exists x_n (y_1, \ldots, y_m \in \mathrm{M} \wedge z = X \wedge \varphi)\} \urcorner$$

is regular.

If X is arbitrary, then

$$\ulcorner \{z | z \in \mathrm{M} \wedge \exists x_1 \ldots \exists x_n (y_1, \ldots, y_m \in \mathrm{M} \wedge z = X \wedge \varphi)\} \urcorner$$

is regular. But this is equal to

$$\ulcorner \{z | \exists x_1 \ldots \exists x_n (y_1, \ldots, y_m, X \in \mathrm{M} \wedge z = X \wedge \varphi)\} \urcorner.$$

So we have the following theorem:

(97) *Let X be regular of signature σ and φ be regular of signature τ. Let $y_1, \ldots, y_m$ be the M-variables of the superposition of σ and τ which are among $x_1, \ldots, x_n$ and let $z_1, \ldots, z_l$ be the remaining ones. If X is a E-term, then*

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M} \urcorner \vdash_2 \ulcorner \{X |_{x_1, \ldots, x_n} y_1, \ldots, y_m \in \mathrm{M} \wedge \varphi\} \in \mathfrak{V} \urcorner.$$

If X is arbitrary, then

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M} \urcorner \vdash \ulcorner \{X |_{x_1, \ldots, x_n} y_1, \ldots, y_m, X \in \mathrm{M} \wedge \varphi\} \in \mathfrak{V} \urcorner.$$

Let now X be the n-tupel term $\ulcorner\langle x_1, \ldots, x_n\rangle\urcorner$ which is of signature $\langle 0 | 0, \ldots, 0\rangle$. The M-variables of $\ulcorner z = X \wedge \varphi\urcorner$ are then the free variables of φ which are not E-variables and among $x_1, \ldots, x_n$, and the M-variables of φ which are not among $x_1, \ldots, x_n$. Therefore we have:

(98) *Let φ be regular; let $y_1, \ldots, y_m$ be the free variables of φ which are not* E*-variables and are among $x_1, \ldots, x_n$, and let $z_1, \ldots, z_l$ be the* M*-variables of φ which are not among $x_1, \ldots, x_n$. Then:*

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\{x_1, \ldots, x_n | y_1, \ldots, y_m \in \mathrm{M} \wedge \varphi\} \in \mho\urcorner.$$

A special case for $n = 2$ and $m = 0$:

(99) *Let φ be regular, u, v* E*-variables of φ and $z_1, \ldots, z_l$ the* M*-variables of φ. Then:*

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\{u, v | \varphi\} \in \mho\urcorner.$$

Conditions φ of this sort, often without any M-variables, are common in relation theory.

We now consider what relativizations are needed to make

$$\ulcorner\{w, x_1, \ldots, x_n | w = X \wedge \varphi\}\urcorner$$

real. Let $y_1, \ldots, y_m$ be the free variables of $\ulcorner w = X \wedge \varphi\urcorner$ which are not E-variables and are among $x_1, \ldots, x_n$, and let $z_1, \ldots, z_l$ be the M-variables of $\ulcorner w = X \wedge \varphi\urcorner$, which are not among $x_1, \ldots, x_n$. If X is a E-term, then w is a E-variable, and according to (98)

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\{w, x_1, \ldots, x_n | y_1, \ldots, y_m \in \mathrm{M} \wedge w = X \wedge \varphi\} \in \mho\urcorner.$$

If X is arbitrary, then we can consider w a M-variable, and we have:

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\{w, x_1, \ldots, x_n | y_1, \ldots, y_m, w \in \mathrm{M} \wedge w = X \wedge \varphi\} \in \mho\urcorner.$$

But $\ulcorner w \in \mathrm{M} \wedge w = X\urcorner$ can be replaced by $\ulcorner X \in \mathrm{M} \wedge w = X\urcorner$.
So we have the following theorem:

(100) *Let X be regular of signature σ and φ be regular of signature τ. Let $y_1, \ldots, y_m$ be the variables of the superposition of σ and τ that are not* E*-variables and are among $x_1, \ldots, x_n$ and let $z_1, \ldots, z_l$ be the remaining* M*-variables.*

If X is a E*-term, then:*

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\langle x_1, \ldots, x_n \mapsto X | y_1, \ldots, y_m \in \mathrm{M} \wedge \varphi\rangle \in \mho\urcorner.$$

If X is arbitrary, then:

$$\ulcorner z_1, \ldots, z_l \in \mathrm{M}\urcorner \vdash_2 \ulcorner\langle x_1, \ldots, x_n \mapsto X | y_1, \ldots, y_m, X \in \mathrm{M} \wedge \varphi\rangle \in \mho\urcorner.$$

We consider some examples.

First we see that the projections are real:

(101) $$\vdash_2 \ulcorner \mathrm{pr}_{n,k} \in \mathfrak{V} \urcorner \quad (\text{for } 1 \leqslant k \leqslant n),$$

$\ulcorner\{x\}\urcorner$ is of signature $\langle 1|0\rangle$, $\ulcorner x \cup y\urcorner$ is of signature $\langle 1|1,1\rangle$. Thus $\ulcorner x \cup \{x\}\urcorner$ is of signature $\langle 1|\mathrm{M}\rangle$. Therefore we have by (100):

$$\vdash_2 \ulcorner \langle x \mapsto x \cup \{x\} \,|\, x,\, x \cup \{x\} \in \mathrm{M}\rangle \in \mathfrak{V}\urcorner.$$

We can drop condition $\ulcorner x \cup \{x\} \in \mathrm{M}\urcorner$, which follows from $\ulcorner x \in \mathrm{M}\urcorner$. Therefore:

(102) $$\vdash_2 \ulcorner \langle x \mapsto x \cup \{x\} \,|\, x \in \mathrm{M}\rangle \in \mathfrak{V}\urcorner.$$

There seems to be no way of extending this function to $\mathfrak{V}$.

Note, however, the following result (which we overlooked for some time). $\ulcorner\{x\}\urcorner$ can be considered to be of signature $\langle 0|0\rangle$. Therefore

(103) $$\vdash_2 \ulcorner \langle x \mapsto \{x\}\rangle \in \mathfrak{V}\urcorner.$$

This proves that in **ZF+M-Comp** there is no phenomenon like the Non-Cantorian classes of **NF** (see [26]). For any class there is the class of its unit subclasses and a bijection between these classes. In fact, by (103) there is such a bijection for $\mathfrak{V}$. Because of (83) the restriction of this function to any real class and also the range of this restriction are real.

The preceeding theorems (about regular expressions and terms) were intended to extend the notion of M-formula to the full language. This is convenient if one wants to work *with* the system. We now adopt a different strategy, namely to reduce the simple formulas to still fewer types of quasiatomic formulas. This is convenient if one wants to work *about* the system.

First we note that we can get rid of the descriptive operator and quasiatomic formulas of type $\ulcorner z = \iota x\varphi\urcorner$ (with z not occurring in φ). By (18), spelling out $\overset{1}{\exists}$ and noting $\ulcorner \bot = \emptyset\urcorner$, we get:

(i) $$\vdash_1 \ulcorner z = \iota x\varphi \leftrightarrow \big(\exists y \forall x(x = y \leftrightarrow \varphi) \wedge \forall x(x = z \leftrightarrow \varphi)\big) \vee \big(\neg \exists y \, \forall x(x = y \leftrightarrow \varphi) \wedge z = \emptyset\big)\urcorner.$$

Similarily we can derive:

(ii) $$\vdash_1 \ulcorner z = \iota x(x \in \mathrm{M} \wedge \varphi) \leftrightarrow \Big(\exists y\big(y \in \mathrm{M} \wedge \forall x(x \in \mathrm{M} \to (x = y \leftrightarrow \varphi))\big) \wedge \forall x\big(x \in \mathrm{M} \to (x = z \leftrightarrow \varphi)\big)\Big) \vee \Big(\neg \exists y\big(y \in \mathrm{M} \wedge \forall x(x \in \mathrm{M} \to (x = y \leftrightarrow \varphi))\big) \wedge z = \emptyset\urcorner\Big).$$

Now suppose that $\ulcorner z = \iota x\varphi \urcorner$, $\ulcorner z = \iota x(x \in \mathrm{M} \wedge \varphi)\urcorner$ are simple and usual M-formulas. In the first case x is a E-variable or C-variable of φ, in the second case x is arbitrary. Suppose now that we can logically equivalent transform φ into a simple and usual M-formula ψ of the same signature as φ without description terms. Then we replace φ by ψ in (i) and (ii). Then we give to x, y, z in the shown subformulas $\ulcorner x = y\urcorner$, $\ulcorner x = z\urcorner$ the same index as z has on the left-hand side. Then we replace $\ulcorner z = \emptyset\urcorner$ by $\ulcorner \exists y(y \in \mathrm{M} \wedge y = z \wedge \forall x\, x \notin y)\urcorner$ or by $\ulcorner \forall x\, x \notin z\urcorner$ as z has index 0 or has index 1.

In case (ii) we in addition give index 0 to x and y in the shown subformulas $\ulcorner y \in \mathrm{M}\urcorner$, $\ulcorner x \in \mathrm{M}\urcorner$. If we finally make the bound variables all different, then we see that the right-hand sides become simple and usual M-formulas of the same signature as $\ulcorner z = \iota x\varphi\urcorner$, $\ulcorner z = \iota x(x \in \mathrm{M} \wedge \varphi)\urcorner$ but without description terms. Repeating this, we see that any simple and usual formula is logically equivalent to a simple and usual formula of the same signature but without description terms.

In a similar way we can eliminate quasiatomic formulas of the type $\ulcorner z = \{x|\varphi\}\urcorner$ (z not in φ). We use the equivalences:

(iii) $\quad \vdash_1 \ulcorner z = \{x|\varphi\} \leftrightarrow \forall x(x \in z \leftrightarrow \varphi)\urcorner$ (if x is a E-variable of φ),

(iv) $\quad \vdash_1 \ulcorner z = \{x|x \in \mathrm{M} \wedge \varphi\} \leftrightarrow \forall x(x \in z \rightarrow x \in \mathrm{M}) \wedge \forall x\big(x \in \mathrm{M} \rightarrow (x \in z \leftrightarrow \varphi)\big)\urcorner$
(if x is arbitrary).

We can eliminate quasiatomic formulas of type $\ulcorner z \in \{x|\varphi\}\urcorner$ using the equivalence:

(v) $$\vdash \ulcorner z \in \{x|\varphi\} \leftrightarrow \varphi_z^x\urcorner.$$

In any case it is easy to check that, if we replace φ by ψ (of same signature but without subformulas of these types) and make all bound variables different, the right-hand sides become M-formulas of the same signatures as $\ulcorner z = \{x|\varphi\}\urcorner$, $\ulcorner z = \{x|x \in \mathrm{M} \wedge \varphi\}\urcorner$, $\ulcorner z \in \{x|\varphi\}\urcorner$.

Repeating this procedure we can completely eliminate quasiatomic formulas of the types $\ulcorner z = \{x|\varphi\}\urcorner$, $\ulcorner z \in \{x|\varphi\}\urcorner$.

We formulate our results as a theorem:

(104) *Any* M-*expression is equivalent to a* M-*formula of the same signature containing only the following types of quasiatomic formulas* (*which are actually atomic*): $\ulcorner x \in y\urcorner$, $\ulcorner x \in \mathrm{M}\urcorner$, $\ulcorner x = y\urcorner$, $\ulcorner z = \langle x, y\rangle\urcorner$.

So it would in principle be possible to restrict the notion of simple formula and drop clauses (iv), (v), (vi) in (78). Using an additional axiom, we can go even further.

Suppose that we add (53), i. e. the axiom **OP′**, identifying $\langle x, y\rangle$ with the set $\{\{x\}, \{x, y\}\}$, which is often used as an ordered pair. There is no problem in doing so, since the sets $\{\{x\}, \{x, y\}\}$ have all the properties of ordered pairs and in addition the property given in clause (iii) of (78), namely $\ulcorner z = \{\{x\}, \{x, y\}\}\urcorner$ is a M-expression of signature $\langle 0, 0, 0\rangle$. This was already shown at the end of section 5. So we can identify $\langle x, y\rangle$ and $\{\{x\}, \{x, y\}\}$.

(105) *It is relative consistent to add axiom* **OP′** *to* **ZF+M-Comp**.

Note that the formula given in section 5 was simple and without descriptive operator and classifier. Adopting axiom **OP′**, we can therefore eliminate the primitive notion $\langle \ldots, \ldots\rangle$ from M-expressions without changing signatures.

Finally we can also eliminate the equality symbol in M-expressions.

We can use the equivalences:

$$\vdash_1 \ulcorner x = y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y)\urcorner$$

(if $\ulcorner x = y\urcorner$ is considered to be of signature $\langle 1, 1\rangle$) and

$$\vdash_1 \ulcorner x = y \leftrightarrow \forall z(x \in z \leftrightarrow y \in z)\urcorner$$

(if $\ulcorner x = y\urcorner$ is considered to be of signature $\langle 0, 0\rangle$).

So we finally see:

(106) *Adopting axiom* **OP′**, *any* M-*expression is equivalent to a* M-*formula containing only the following types of quasiatomic formulas*:

$$\ulcorner x \in y\urcorner, \ulcorner x \in \mathrm{M}\urcorner.$$

For formulas of this type, in definition (78) only clause (i) remains. But then the M-variables are simply the variables occurring on both sides of the $\in$.

So let us consider the following rather simple comprehension principle.

(107) **M-Comp***:

Let φ be usual and only contain connectives, quantifiers and atomic formulas of type $\ulcorner x \in y\urcorner$, $\ulcorner x \in \mathrm{M}\urcorner$. Let z be different from x and not in φ. If we relativize in the formula

$$\ulcorner\forall \ldots \exists z \forall x(x \in z \leftrightarrow \varphi)\urcorner$$

all variables occurring on both sides of $\in$ to M, *then the result is an axiom.*

Let **ZF+M-Comp*** be the theory which has as axioms the axioms of **ZF+Cls**, and **M-Comp***. We have thus seen that **ZF+M-Comp*+OP′** is the same as **ZF+M-Comp+OP′**.

The theory **ZF+M-Comp*** (restricted to a language without primitive ordered pairs) is actually the theory **G*** considered in [14]. We did not use the indexing there which is not needed for the formulation of **M-Comp***. There seems to be one more difference. In [14] we were writing $\ulcorner \mathrm{M}x \urcorner$ instead of $\ulcorner x \in \mathrm{M} \urcorner$ and accordingly such x did not bring in additional left-of-ϵ or right-of-ϵ occurrences of x. But we can also avoid additional such occurrences of x with our type of atomic formulas. If the x in $\ulcorner \mathrm{M}x \urcorner$ is not to be counted as a right-of-ϵ occurrence, then we translate it by $\ulcorner x \in \mathrm{M} \urcorner$. If the x in $\ulcorner \mathrm{M}x \urcorner$ is not to be counted as a left-of-ϵ occurrence, then we translate it by $\ulcorner \exists y (y \in \mathrm{M} \wedge \forall z (z \in y \leftrightarrow z \in x)) \urcorner$. The equivalence of the systems **G*** of [14] and **ZF+M-Comp*** is thus completely settled.

In clause (iii) of (78) we say that $\ulcorner z = \langle x, y \rangle \urcorner$ is of signature $\langle 0, 0, 0 \rangle$. Our next problem will be the question whether we can consider it also of signature $\langle 1, 1, 1 \rangle$. Actually, this is impossible in **ZF+M-Comp+OP′**. We first note the following theorem which is already a theorem of **ZF+Cls**:

(108) $$\vdash_1 \ulcorner \{\langle x, x \rangle \mid \langle x, x \rangle \notin x\} \notin \mathcal{V} \urcorner.$$

For the proof let X be $\ulcorner \{\langle x, x \rangle \mid \langle x, x \rangle \notin x\} \urcorner$. Then for any y:

$$\ulcorner \langle y, y \rangle \in \mathcal{V} \urcorner \vdash \ulcorner \langle y, y \rangle \in X \leftrightarrow \exists x (\langle y, y \rangle = \langle x, x \rangle \wedge \langle x, x \rangle \notin x) \urcorner.$$

The premise $\ulcorner \langle y, y \rangle \in \mathcal{V} \urcorner$ is provable. Furthermore $\ulcorner \langle y, y \rangle = \langle x, x \rangle \urcorner$ is equivalent to $\ulcorner y = x \urcorner$, and then the right-hand side of the equivalence is equivalent to $\ulcorner \langle y, y \rangle \notin y \urcorner$. So we have:

$$\vdash_1 \ulcorner \langle y, y \rangle \in X \leftrightarrow \langle y, y \rangle \notin y \urcorner.$$

Substituting X:

$$\ulcorner X \in \mathcal{V} \urcorner \vdash_1 \ulcorner \langle X, X \rangle \in X \leftrightarrow \langle X, X \rangle \notin X \urcorner.$$

This gives $\vdash_1 \ulcorner X \notin \mathcal{V} \urcorner$.

Suppose now in **ZF+M-Comp+OP′** $\ulcorner z = \langle x, y \rangle \urcorner$ is of signature $\langle 1, 1, 1 \rangle$. Then $\ulcorner \langle x, x \rangle \urcorner$ is a C-term of signature $\langle 1 | 1 \rangle$. Since by axiom **OP′** we get $\ulcorner \langle x, x \rangle \in \mathrm{M} \urcorner$, it is also an E-term and of signature $\langle 0 | 1 \rangle$. Substituting this for y in formula $\ulcorner y \notin x \urcorner$ (of signature $\langle 0, 1 \rangle$) one gets formula $\ulcorner \langle x, x \rangle \notin x \urcorner$ of signature $\langle 1 \rangle$.

Therefore by (97) we could derive:

$$\ulcorner \{\langle x, x \rangle \mid \langle x, x \rangle \in \mathrm{M} \wedge \langle x, x \rangle \notin x\} \in \mathcal{V} \urcorner.$$

Again by **OP′** we could derive:

$$\ulcorner \langle x, x \rangle \in \mathrm{M} \wedge \langle x, x \rangle \notin x \leftrightarrow \langle x, x \rangle \notin x \urcorner.$$

Therefore we could derive:

$$\ulcorner \{\langle x, x \rangle \mid \langle x, x \rangle \notin x\} \in \mathcal{V} \urcorner,$$

contradicting (108).

However, in our proof we used that all ordered pairs are sets.

The question arises whether we can have $\ulcorner z = \langle x, y\rangle \urcorner$ of signature $\langle 1, 1, 1\rangle$ if we drop this assumption.

This is indeed possible, but we have to take another ordered pair. Let us try the Schmidt pair which is a set iff both components are sets. It is easily seen that

$$\ulcorner z = \big\{\{\{u\}\} \,|\, u \in x\big\} \cup \big\{\{\emptyset, \{v\}\} \,|\, v \in y\big\}\urcorner$$

is a M-expression of signature $\langle 1, 1, 1\rangle$. The problem is now whether we can have it also of signature $\langle 0, 0, 0\rangle$. This signature is quite indispensable whereas signature $\langle 1, 1, 1\rangle$ would only be a welcome addition.

In fact, we can have both signatures.

(109) *It is relative consistent to add* **OP″** *to* **ZF+M-Comp** *and replace clause* (iii) *in* (78) *by*

(iii′) *the indicated occurrences of z, x and y in a subformula $\ulcorner z = \langle x, y\rangle \urcorner$ get all the same index.*

In order to see this, we repeat in a modified form the consistency proof of [14].

The consistency proof is relative to a set theory which is not "over classes" and therefore intuitively better accessable. It is a set theory with urelements, set over urelements, and classes of such sets which are all ultimate if proper. Of course, the set axioms have to be changed a little bit. Since we want to admit urelements, we drop the extensionality axiom. Since we will also have ultimate classes, we take the Kühnrich pairs, i. e. adopt **OP‴** (29). We must modify **A 1** which contradicts ultimate classes. We replace it by

(110) **A 1′**: $\ulcorner x, y \in \mathrm{EL} \to \{x, y\} \in \mathrm{M}\urcorner$.

The replacement schema is now claimed only for functions mapping elements into elements.

(111) **S 5′**: $\ulcorner x \in \mathrm{M} \wedge (\forall y \in x)(\overset{1}{\exists} z \in \mathrm{EL})\varphi \to \{z \,|\, z \in \mathrm{EL} \wedge (\exists y \in x)\varphi\} \in \mathrm{M}\urcorner$.

Class existence axioms are given by the so-called impredicative comprehension principle saying that any class of elements is real which first occurred in Quine's system **ML** [21] and was first brought into connection to set theory by Wang [31].

(112) **I-Comp**: $\ulcorner X \subseteq \mathrm{EL} \to X \in \mathfrak{V}\urcorner$.

By this comprehension principle UR, M, V, On turn out to be proper classes, but UC (which is equal to PC), Cls, $\mathfrak{V}$ are virtual.

---

(29) Ordered pairs for ultimate classes are not urgently needed and one often gets along without them. But in view of our general frame **CRF** (section 3) it would be odd to have real objects without ordered pairs.

Since all proper classes are ultimate, any class being an element must be a set. Conversely of course sets should be elements and not ultimate. Therefore our last axiom is:

(113) $\ulcorner \mathrm{M} = \mathrm{EL} \cap \mathrm{Cls}\urcorner$.

The primitive notions M and $\langle\ ,\ \rangle$ could in principle be eliminated, and this is usually done in set theory. Let us call this set theory $\mathbf{ZFU}^{(2)}$, the second order Zermelo–Frankel theory with urelements admitted(30). This type of set theory is well-known and widely used. Therefore a relative consistency proof to $\mathbf{ZFU}^{(2)}$ can claim to give some plausibility to the system under consideration. This is still true if we add other rather strong but familiar axioms.

As additional axiom we take:

(114) $\ulcorner$*$\varkappa$ is a strongly inaccessible cardinal and* UR *is a set of cardinality $\varkappa$*$\urcorner$.

We now use this set theory as metalanguage. We want to get a model Δ such that ϵ^{Δ} is the natural element relation, UR^{Δ}, M^{Δ}, UC^{Δ} have all the same cardinality $\varkappa$ (so that a bijection F between UC^{Δ} and UR^{Δ} is possible) and which is "full" in sense that M^{Δ} is closed under the formation of images by functions mapping elements of M^{Δ} onto subclasses of EL^{Δ} (this will give the Ersetzungsschema)(31).

Moreover, the formation of ordered pairs of elements and classes should "run parallel" with respect to the bijection F.

First we consider a model Δ'' such that $\epsilon^{\Delta''}$, $\langle\ ,\ \rangle^{\Delta''}$ are natural (i. e. the same as in $\mathbf{ZFU}^{(2)}$ plus axiom (114)) and:

$\mathrm{UR}^{\Delta''} = \mathrm{UR}$ = set of urelements.

$\mathrm{M}^{\Delta''} = \mathrm{H}_{\varkappa}$ = set of sets that are hereditarily of cardinality less than $\varkappa$.

$\mathrm{UC}^{\Delta''} = \mathrm{P}(\mathrm{UR} \cup \mathrm{H}_{\varkappa}) \setminus \mathrm{H}_{\varkappa}$ = set of sets of cardinality $\varkappa$ that contain only urelements and sets of $\mathrm{H}_{\varkappa}$.

$\mathfrak{V}^{\Delta''}$, the individual domain of Δ, is the union of these three disjoint sets.

Δ'' is a model of $\mathbf{ZFU}^{(2)}$ such that $\mathrm{UR}^{\Delta''}\mathrm{M}^{\Delta''}$ have the same cardinality $\varkappa$, $\mathrm{UC}^{\Delta''}$ has greater cardinality. Using the downward Löwenheim–Skolem theorem, we thin $\mathrm{UC}^{\Delta''}$ down to cardinality $\varkappa$ leaving $\mathrm{UR}^{\Delta''}$, $\mathrm{M}^{\Delta''}$ unchanged. So we get a model Δ' of $\mathbf{ZFU}^{(2)}$ such that $\epsilon^{\Delta'}$, $\langle x, y\rangle^{\Delta'}$ are natural and

$$\mathrm{UR}^{\Delta'} = \mathrm{UR}^{\Delta''}, \quad \mathrm{M}^{\Delta'} = \mathrm{M}^{\Delta''}, \quad \mathrm{UC}^{\Delta'} \subseteq \mathrm{UC}^{\Delta''},$$

(30) We use the term "second order" since we look at this theory in a way that the class variables are considered as second order variables of the element variables, c. f. footnote (15).

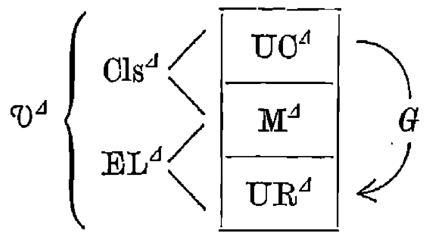
(31) In [14] we used only a denumerable model that was not full. The improvement was suggested by R. B. Jensen.

$\mathrm{UC}^{\Delta'}$ has cardinality $\varkappa$. Moreover, Δ' is full (in the sense given above) since Δ'' is full and $\mathrm{UC}^{\Delta''}$, $\mathrm{M}^{\Delta''}$ have not changed.

The model Δ comes from Δ' by redefining the ordered pair and leaving everything else unchanged. So we simply omit the primes and define $\langle x, y\rangle^{\Delta}$ new for $x, y \in \mathcal{V}^{\Delta}$. This definition will come later. We first describe the model Δ^*. Since UC^{Δ} and UR^{Δ} have the same cardinality, there is a function F which is a one-one mapping of UC^{Δ} onto UR^{Δ}. We extend F to a function G by putting

$$G = F \cup \mathrm{id}_{\mathrm{M}^{\Delta}}.$$

So G maps $\mathrm{UC}^{\Delta} \cup \mathrm{M}^{\Delta}$ (which is Cls^{Δ}) one-one onto $\mathrm{UR}^{\Delta} \cup \mathrm{M}^{\Delta}$ (which is EL^{Δ}), and G is the identity on M^{Δ}. The situation is visualized in a diagram:



Model Δ

We use the function G to replace the urelements by the ultimate classes and define a new model Δ^*. The domain $\mathcal{V}^{\Delta^*}$ of Δ^* is Cls^{Δ} and M^{Δ^*} is M^{Δ}. We redefine the ϵ relation as follows:

$$x \epsilon^{\Delta^*} y \leftrightarrow G(x) \epsilon^{\Delta} y \qquad (\text{for } x, y \in \mathrm{Cls}^{\Delta}).$$

For the ordered pairs we could take the Kühnrich pairs restricted to classes. But of course the Schmidt pairs are simpler([32]). So we put:

$$\langle x, y\rangle^{\Delta^*} = \Big(\big\{\{\{u\}\} \mid u \in x\big\} \cup \big\{\{\emptyset, \{v\}\} \mid v \in y\big\}\Big)^{\Delta^*}.$$

This ends the description of the model Δ^*. We now return to Δ and define the ordered pairs there: we put

$$\langle x, y\rangle^{\Delta} = \langle x, y\rangle^{\Delta^*} \qquad \text{if } x, y \in \mathrm{Cls}^{\Delta},$$
$$\langle x, y\rangle^{\Delta} = G\big(\langle G^{-1}(x), G^{-1}(y)\rangle^{\Delta^*}\big) \qquad \text{if } x, y \in \mathrm{EL}^{\Delta}.$$

If $x, y \in \mathrm{M}^{\Delta}$, then both lines define the same value. The pair is not yet defined for $x \in \mathrm{UC}^{\Delta}$ and $y \in \mathrm{UR}^{\Delta}$ or for $x \in \mathrm{UR}^{\Delta}$ and $y \in \mathrm{UR}^{\Delta}$. For these arguments we put:

$$\langle x, y\rangle^{\Delta} = H(x, y),$$

([32]) Any two-place function will do if it maps $\mathrm{M}^{\Delta} \times \mathrm{M}^{\Delta}$ one-one into M^{Δ} and $(\mathrm{Cls}^{\Delta} \times \mathrm{Cls}^{\Delta}) \setminus (\mathrm{M}^{\Delta} \times \mathrm{M}^{\Delta})$ one-one into UC^{Δ}.

where H is some one-one function mapping $(\mathrm{UC}^\Delta \times \mathrm{UR}^\Delta) \cup (\mathrm{UR}^\Delta \times \mathrm{UC}^\Delta)$ in M^Δ such that the values are not pairs already defined. This is a suitable definition of ordered pairs. The axiom of ordered pairs e. g. is fulfilled since it is a one-one function. For the one-one property it is important that $\langle x, y\rangle^{\Delta^*} \notin \mathrm{M}^\Delta$ if one of x or y is in UC^Δ, since otherwise $\langle x, y\rangle$ and $\langle G(x), G(y)\rangle$ would be different but have the same ordered pair in Δ.

The axioms of $\mathbf{ZFU}^{(2)}$ are actually given (with the exception of $\mathbf{OP}'''$) without mentioning ordered pairs. So Δ will be a model of $\mathbf{ZFU}^{(2)}$ except $\mathbf{OP}'''$. (Note that in the schema $\mathbf{S\,5}'$ the formula φ is arbitrary, so identifying $\langle x, y\rangle$ with some other one-one function will make no difference.)

We now give a syntactical definition which parallels (78):

(115) *Let φ be usual and simple and assume that indices are given to the variable occurrences in φ according to* (78) *but with* (iii) *replaced by* (iii′) *of* (109).
Then $\varphi^{\mathrm{E,M,C}}$ comes from φ by relativizing to EL *all variables getting index* 0 *throughout φ, relativizing to* M *all variables getting both indices in φ, relativizing to* Cls *all variables getting index* 1 *throughout φ. We call $\varphi^{\mathrm{E,M,C}}$ a* E-M-C-*formula.*

Let us speak of signatures of E-M-C-formulas in the same way as we spoke of signatures of M-formulas. The signatures indicate what indices are given to the free variables of the formula.

Now consider two assignments h, h^* of values from $\mathcal{V}^\Delta$, $\mathcal{V}^{\Delta^*}$ to the variables of some signature σ. We say that h, h^* *coherent* with respect to σ if

$$h(x) = h^*(x) \qquad \text{if } x \text{ is a C-variable of } \sigma,$$

$$h(x) = G(h^*(x)) \qquad \text{if } x \text{ is a E-variable of } \sigma,$$

$$h(x) = h^*(x) = G(h^*(x)) \in \mathrm{M}^\Delta \qquad \text{if } x \text{ is a M-variable of } \sigma.$$

In particular, $h(x)$ is a class of Δ if x is a C-variable, and an element of Δ if x is a E-variable.

The main lemma is:

(116) *If φ is simple and usual and h, h^* are coherent, then it holds*

$$\Delta \models \varphi^{\mathrm{E,M,C}}[h] \leftrightarrow \Delta^* \models \varphi^{\mathrm{M}}[h^*].$$

The proof is straight forward by induction on the structure of φ. We consider the case that φ and also $\varphi^{\mathrm{E,M,C}}, \varphi^{\mathrm{M}}$ are $\ulcorner z = \langle x, y\rangle\urcorner$. Let $a = h(z)$, $b = h(x)$, $c = h(y)$, $a^* = h^*(z)$, $b^* = h^*(x)$, $c^* = h^*(y)$.

If all variables get index 0, then by coherence

$$a = G(a^*), \quad b = G(b^*), \quad c = G(c^*)$$

and

$$\begin{aligned} a = \langle b, c\rangle^{\Delta} \leftrightarrow a &= G(\langle G^{-1}(b), G^{-1}(c)\rangle^{\Delta^*}) \\ \leftrightarrow G^{-1}(a) &= \langle G^{-1}(b), G^{-1}(c)\rangle^{\Delta^*} \\ \leftrightarrow a^* &= \langle b^*, c^*\rangle^{\Delta^*}. \end{aligned}$$

If all variables get index 1, then by coherence

$$a = a^*, \quad b = b^*, \quad c = c^*,$$

and:

$$a = \langle b, c\rangle^{\Delta} \leftrightarrow a = \langle b, c\rangle^{\Delta^*} \leftrightarrow a^* = \langle b^*, c^*\rangle^{\Delta^*}.$$

In any case we have:

$$\Delta \models \ulcorner z = \langle x, y\rangle\urcorner[h] \leftrightarrow \Delta^* \models \ulcorner z = \langle x, y\rangle\urcorner[h^*].$$

The other steps are as in [14].

For the quantifier case the following is important:

The set of possible values for a quantifier in Δ relativized to EL is EL^{Δ}. This is mapped by G^{-1} one-one onto Cls^{Δ} which is $\mathfrak{V}^{\Delta^*}$, the whole domain of individuals of Δ. The set of possible values for a quantifier in Δ relativized to Cls is Cls^{Δ} which is $\mathfrak{V}^{\Delta^*}$. Because of this, we can drop the relativizations to EL for E-variables and to Cls for C-variables when passing from Δ to Δ^*.

Next one shows that axioms of **ZF + M-Comp** (except **S 4**, **S 5**) can be brought into the from φ^{M} for some $\varphi^{\mathrm{E,M,C}}$ which is immediately equivalent to an axiom of $\mathbf{ZFU}^{(2)}$. This is done in [14]. But **S 5** is also true.

Let f be a function of Δ^* mapping x (for $x \in \mathrm{M}^{\Delta^*}$) onto a class y ($y \in \mathrm{Cls}^{\Delta^*}$). Then $G^{-1} \circ f \circ G$ is a function of Δ mapping x onto y. Since $x \in \mathrm{M}^{\Delta}$ and Δ is full $y \in \mathrm{M}^{\Delta} = \mathrm{M}^{\Delta^*}$.

Finally by the construction of $\langle\ ,\ \rangle^{\Delta^*}$, the axiom **OP″** holds in Δ^*. This ends the proof of (109).

Finally we give a brief remark about possible extensions:

In our system **ZF + M-Comp** we excluded urelements. We now see how we could get a similar system **ZFU + M-Comp** with urelements admitted. The function F is then supposed only to be into UR^{Δ} and not onto UR^{Δ}. So there may be some urelements left in the resulting model Δ^*.

However, then in the main lemma we could not drop the relativizations to Cls, since a quantifier over the classes in Δ would not correspond to a quantifier over the universe in Δ^*. So we would have to use a notion $\varphi^{\mathrm{M,C}}$ rather than φ^{M} in the comprehension schema.

7. A combination of set theory and stratification

We now want to add to **ZF+Cls** the class existence schema of Quine's system **NF** [20], viz. the stratification schema([33]). A stratified formula can approximately be described as a formula of a one-sorted class theoretical language which is "like a type theoretical formula", i. e. which by putting superscripts to the variables can be transformed into a formula of simple type theory. In **NF** any stratified condition defines a real class. **NF** has been used as a basis of mathematics (e. g. see [26]). However, **NF** has some odd features that make it — as we think — hardly acceptable as a working set theory. E. g. the axiom of choice and the Aussonderungsschema fail. A minor point is that **NF** needs additional axioms: one cannot prove that the class of natural numbers is real (unless **NF** is inconsistent). We think that **NF** is a typical class theory and that the notion of set does not occur in **NF** at all. Since we want to deal with sets too, we need the additional primitive M (**NF** is usually given in a language with only ϵ as primitive). Formulas $x \epsilon \mathrm{M}$ are stratified whatever the index of x is. Furthermore, we treat the ordered pairs as primitive and require that $z = \langle x, y\rangle$ is stratified if z, x, y get all the same index. We say for this that the pair is *homogeneous in type*. This is very convenient for the theory of relations and functions as we will see. Moreover, we think that also intuitively ordered pairs should be homogeneous in type, that all n-place relations and functions over objects of some fixed type should be for all n of the same next higher type, and that n-tuples for all n should be of the same type as the arguments. Most definitions of ordered pairs (e. g. the ones in (53), (54), (55)) are not homogeneous in type. But actually Quine has shown in [22] that there is a definition of homogeneous ordered pairs for **NF**([34]). Unfortunately, the definition is rather complicated. So we consider it simpler to treat the homogeneous ordered pair axiomatically outright. Because of Quine's result we are sure that this is an eliminable extension of **NF**.

The stratification is different from the indexing in section 5 in several respects. First we can allow any non-negative integer as index([35]).

Secondly the indices are given to the variables rather than the variable occurrences (different occurrences of the same variable in an usual formula get always the same index).

Thirdly there is no flexibility in changing the signature (if one has proved that something is a set) as is in **ZF+M-Comp**. Therefore the indexing

---

([33]) A combination of set theory and stratification is also considered in [3] and [9]. See section 9 (historic remark).

([34]) Rosser uses in his book [26] the Quine pairs.

([35]) Of course, introducing negative indices would make no difference for **NF**.

of the variables can be continued in a unique way to an indexing of all general terms in a stratified formula. This makes it easy to define the notion of stratification not only for simple formulas but for arbitrary formulas and also for terms.

For simplicity we give the definition only for usual formulas and terms and require that any alphabetic variant of a stratified formula or term is also stratified.

(117) *A usual formula or usual general term is stratified iff there is an assignment* $\langle X \mapsto \operatorname{ind}(X)\rangle$ *of indices from the non negative integers to the subterms (other than* M) *such that:*

(i) *if* $\ulcorner X \in Y \urcorner$ *is a subexpression, then* $\operatorname{ind}(X)+1 = \operatorname{ind}(Y)$,

(ii) *if* $\ulcorner X = Y \urcorner$ *is a subexpression, then* $\operatorname{ind}(X) = \operatorname{ind}(Y)$,

(iii) *if* $\ulcorner\langle X, Y\rangle\urcorner$ *is a subterm, then* $\operatorname{ind}(\ulcorner\langle X, Y\rangle\urcorner) = \operatorname{ind}(X) = \operatorname{ind}(Y)$,

(iv) *if* $\ulcorner \iota x\varphi \urcorner$ *is a subterm, then* $\operatorname{ind}(\ulcorner \iota x\varphi\urcorner) = \operatorname{ind} x$,

(v) *if* $\ulcorner\{x \mid \varphi\}\urcorner$ *is a subterm, then* $\operatorname{ind}(\ulcorner\{x\mid\varphi\}\urcorner) = \operatorname{ind} x+1$,

(vi) M *may get different positive indices at different occurrenes.*

The last condition may look strange at first sight. But it is quite natural since other constant terms are also "floating in type". E. g. if $\mho$ occurs in a formula φ at different places than, if φ is made usual, we have to use different bound variables, say: $\ulcorner\{x \mid x = x\}\urcorner$, $\ulcorner\{y \mid y = y\}\urcorner$. These terms can get quite different positive indices.

The stratification schema is:

(118) **Strat**: *Let* z *not occur in* $\ulcorner\{u \mid \varphi\}\urcorner$. *Then*

$$\ulcorner \exists z\, z = \{u \mid \varphi\}\urcorner$$

is an axiom if it is stratified.

NF is the theory with axioms **Ext**, **Strat.**

ZF+NF is the theory which has the axioms of **ZF+Cls** and all instances of **Strat.**

We write $\Sigma \vdash_3 \varphi$ to indicate $\Sigma \cup A \vdash \varphi$ for some set A of (closed) formulas which are axioms of **ZF+NF**. Thus $\vdash_3 \varphi$ means that φ is a theorem of **ZF+NF.**

ZF+NF seems to be (if consistent) a more satisfactory basis for mathematics than **NF** alone. Since it contains a good set theory, the usual set theoretical reasoning is possible. All theorems of **ZF+Cls** are also theorems of **ZF+NF.** Moreover, we have a lot of proper classes in **ZF+NF.** The fact that these proper classes have some oddities mentioned above is perhaps excusable. One expects that proper classes do not have all the nice properties of sets.

In order to show the strength of **ZF+NF**, we present some applications. We say nothing about the set theoretical part which is assumed to be covered by section 4. All applications are therefore already theorems of **NF** alone. Since **NF** is well investigated the theorems are not new. But perhaps the presentation may be of some interest.

Of course, any stratified term denotes a real class. Let X have under some indexing the index i. Without loss of generality we can assume $i > 0$ since we can otherwise add 1 to all indices. Take new variables z, u and put $\mathrm{ind}(z) = i$, $\mathrm{ind}(u) = i-1$. Then

$$\ulcorner \exists z\, z = \{u \mid u \in X\} \urcorner$$

is stratified and an axiom. This gives

(119) *If X is stratified, then* $\underset{3}{\vdash} \ulcorner X \in \mathcal{V} \urcorner$.

It is convenient to introduce stratification signatures of formulas and terms.

(120) *Let $x_1, \ldots, x_n$ be different and the free variables of φ and X be in $\{x_1, \ldots, x_n\}$.*
Suppose that φ is stratified under some assignment of indices with $\mathrm{ind}(x_k) = i_k$ $(1 \leqslant k \leqslant n)$. Then we say that φ has stratification signature $\langle i_1, \ldots, i_n \rangle$ with respect to $x_1, \ldots, x_n$.
Suppose that X is stratified under some assignment of indices with $\mathrm{ind}(x_k) = i_n$, $\mathrm{ind}(X) = i$. Then we say that φ has stratification signature $\langle i \mid i_1, \ldots, i_n \rangle$ with respect to $x_1, \ldots, x_n$.

Of course, signatures are not unique. When we add to all components the same positive integer, we get a new signature. In the following we give the least signature, and it is always understood with respect to the variables in the order of first occurrence in the given abbreviation.

We will not formulate rules for the composition of stratified formulas and terms since these are quite clear. We only mention that when we superpose signatures, they must match, i. e. may not assign different indices to the same variable. Of course, sometimes signatures can be made matching by adding a positive integer to one of them.

We now give a list of stratified formulas and terms with signatures. One will remark, that most of them did already occur with the same signature in **ZF+M-Comp**. Only $\bigcup x$, $\bigcap x$, $\mathrm{P}x$ have other signatures here, and $\mathbb{U}x$, $\mathbb{\Pi}x$, $\mathbb{P}x$ which are not regular (in the sense of section 6) are new here.

Stratified formulas:

$\ulcorner x \in y \urcorner$	*Signature*: $\langle 0, 1 \rangle$,
$\ulcorner x = y \urcorner$	*Signature*: $\langle 0, 0 \rangle$,

$\ulcorner x \subseteq y \urcorner$	*Signature*: $\langle 1, 1 \rangle$,
$\ulcorner xzy \urcorner$	*Signature*: $\langle 0, 1, 0 \rangle$,
$\ulcorner \mathrm{Relation}(x) \urcorner$, $\ulcorner \mathrm{Function}(x) \urcorner$	*Signature*: $\langle 1 \rangle$,
$\ulcorner z \mid x \to y \urcorner$	*Signature*: $\langle 1, 1, 1 \rangle$,
Stratified terms:	
x	*Signature*: $\langle 0 \mid 0 \rangle$,
$\ulcorner \langle x, y \rangle \urcorner$	*Signature*: $\langle 0 \mid 0, 0 \rangle$,
$\mathrm{M}, \varnothing, \mho$	*Signature*: $\langle 1 \mid \rangle$,
$\ulcorner \{x, y\} \urcorner$	*Signature*: $\langle 1 \mid 0, 0 \rangle$,
$\ulcorner x \cap y \urcorner$, $\ulcorner x \cup y \urcorner$, $\ulcorner x \times y \urcorner$, $\ulcorner x \circ y \urcorner$, $\ulcorner x \restriction y \urcorner$	*Signature*: $\langle 1 \mid 1, 1 \rangle$,
$\ulcorner -x \urcorner$, $\ulcorner \mathrm{id}_x \urcorner$, $\ulcorner x^{-1} \urcorner$, $\ulcorner \mathrm{dom}\, x \urcorner$, $\ulcorner \mathrm{ran}\, x \urcorner$	*Signature*: $\langle 1 \mid 1 \rangle$,
$\ulcorner x(y) \urcorner$	*Signature*: $\langle 0 \mid 1, 0 \rangle$,
$\ulcorner \bigcup x \urcorner$, $\ulcorner \bigcap x \urcorner$, $\ulcorner \mathbb{U} x \urcorner$, $\ulcorner \mathbb{\Pi} x \urcorner$	*Signature*: $\langle 1 \mid 2 \rangle$,
$\ulcorner \mathrm{P} x \urcorner$, $\ulcorner \mathbb{P} x \urcorner$	*Signature*: $\langle 2 \mid 1 \rangle$.

We immediately get the theorem that the following abstraction terms are real:

(121) $\vdash_3 \ulcorner \mathrm{M}, \varnothing, \mho, \{x, y\}, x \cap y, x \cup y, x \times y, x \circ y,$

$x \restriction y,\ -x, \mathrm{id}_x, x^{-1}, \mathrm{dom}\, x, \mathrm{ran}\, x,$

$\bigcup x, \bigcap x, \mathbb{U} x, \mathbb{\Pi} x, \mathrm{P} x, \mathbb{P} x \in \mho \urcorner.$

But we see no way to prove $\ulcorner \mathrm{V} \in \mho \urcorner$, or $\ulcorner \mathrm{On} \in \mho \urcorner$ since the formulation of transitivity is not stratified.

We now give three theorems about generalized abstraction, about relational and functional abstraction. The proofs are immediate if one eliminates the definitions.

(122) *Suppose φ and X are stratified with matching signatures. Then*:

$$\vdash_3 \ulcorner \{X \mid_{x_1 \ldots x_n} \varphi\} \in \mho \urcorner.$$

(123) *Suppose φ is stratified of signature $\langle i, i, \ldots, i, j_1, \ldots, j_m \rangle$ with respect to $x_1, \ldots, x_n, y_1, \ldots, y_m$. Then*:

$$\vdash_3 \ulcorner \{x_1, \ldots, x_n \mid \varphi\} \in \mho \urcorner.$$

(124) *Suppose φ is stratified of signature $\langle i, i, \ldots, i, j_1, \ldots, j_m \rangle$ with respect to $x_1, \ldots, x_n, y_1, \ldots, y_m$ and X is stratified of signature $\langle i \mid i, \ldots, i, k_1, \ldots, k_l \rangle$ with respect to $x_1, \ldots, x_n, z_1, \ldots, z_l$ and both signatures match. Then*:

$$\vdash_3 \ulcorner \langle x_1, \ldots, x_n \mapsto X \mid \varphi \rangle \in \mho \urcorner.$$

In the same way as in **ZF+M-Comp** we see that the projections are real:

(125) $$\vdash_3 \ulcorner \mathrm{pr}_{n,k} \in \mathfrak{V} \urcorner \quad (\text{for } 1 \leqslant k \leqslant n).$$

But for the other examples following (101) we get different results.

Since the term $\ulcorner x \cup \{x\} \urcorner$ is not stratified, we see no possibility to prove the reality of the function occurring in (102). Since $\ulcorner \{x\} \urcorner$ is not of signature $\langle i \,|\, i \rangle$, we cannot prove the reality of the function occurring in (103). Actually, it is well known that one can prove in **NF** that this function is *not* real.

(126) $$\vdash_3 \ulcorner \langle x \mapsto \{x\} \rangle \notin \mathfrak{V} \urcorner.$$

In order to see this, we make the assumption:

$$\ulcorner z \,|\, \{\{x\} \,|\, x \in \mathfrak{V}\} \xrightarrow{\text{onto}} \mathfrak{V} \urcorner.$$

Since $\ulcorner z(\{x\}) \urcorner$ is stratified of signature $\langle 1 \,|\, 2, 0 \rangle$, also $\ulcorner x \notin z(\{x\}) \urcorner$ is stratified of signature $\langle 0, 2 \rangle$. Therefore $\ulcorner \{x \,|\, x \notin z(\{x\})\} \urcorner$ is real, and we can assume:

$$\ulcorner y_0 = \{x \,|\, x \notin z(\{x\})\} \in \mathfrak{V} \urcorner.$$

Under the assumption there is a x_0 such that $y_0 = z(\{x_0\})$. So we have:

$$\ulcorner x \in z(\{x_0\}) \leftrightarrow x \notin z(\{x\}) \urcorner.$$

This gives a contradiction putting x_0 for x. This proves

$$\vdash_3 \ulcorner \neg \exists z\, z \,|\, \{\{x\} \,|\, x \in \mathfrak{V}\} \xrightarrow{\text{onto}} \mathfrak{V} \urcorner.$$

But if $\ulcorner w = \langle x \mapsto \{x\} \rangle \in \mathfrak{V} \urcorner$, since w is one-one, w^{-1} would be such a function. This proves (126).

So **ZF+M-Comp** and **ZF+NF** turn out to be incompatible.

We now want to see that classes of big structures exist in **ZF+NF**. All first order axiom systems are stratified when the relation and function symbols get index 1 and the individual symbols get index 0. Take e. g. the formula $\ulcorner \mathrm{group}(x, y) \urcorner$ saying that x is a class, y a binary composition on x satisfying the group axioms. Let $\ulcorner \mathrm{Group}(z) \urcorner$ be the formula $\ulcorner \exists x \exists y (z = \langle x, y \rangle \wedge \mathrm{group}(x, y)) \urcorner$. Then $\ulcorner \mathrm{group}(x, y) \urcorner$ is stratified with signature $\langle 1, 1 \rangle$ and $\ulcorner \mathrm{Group}(z) \urcorner$ is stratified with signature $\langle 1 \rangle$. Therefore we have:

(127) $$\vdash_3 \ulcorner \{z \,|\, \mathrm{Group}(z)\} \in \mathfrak{V} \urcorner.$$

So the class of all groups (big or small) exists in **ZF+NF**. In a similar way one can see that various other classes of models exist in **ZF+NF**. In **ZF+M-Comp** (using ordered pairs with signature $\langle 1 \,|\, 1, 1 \rangle$) the formula $\ulcorner \mathrm{Group}(z) \urcorner$ would also be regular of signature $\langle 1 \rangle$. But nevertheless (127) would not hold since z is to be relativized to M. So we would only have that the class of all small groups (which have a *set* as individual domain) exists in **ZF+M-Comp**. Of course, this class exists in **ZF+NF** too.

Our next example is the permutation group of a class.

Let $\ulcorner z|x \xrightarrow[\text{onto}]{1-1} x\urcorner$ be $\ulcorner z|x \to x \wedge z^{-1}|x \to x\urcorner$. This is stratified with signature $\langle 1, 1\rangle$.

Let $\ulcorner X_1(x)\urcorner$ be $\ulcorner\{z|z|x \xrightarrow[\text{onto}]{1-1} x\}\urcorner$. This is stratified with signature $\langle 2|1\rangle$. The formula $\ulcorner u, v, w \in X_1(x) \wedge u = v \circ w\urcorner$ is stratified with signature $\langle 1, 1, 1, 1\rangle$. Let $\ulcorner X_2(x)\urcorner$ be $\ulcorner\{u, v, w|u, v, w \in X_1(x) \wedge u = v \circ w\}\urcorner$. This is stratified with signature $\langle 2|1\rangle$. Let

$$\ulcorner\mathrm{Perm}(x)\urcorner \quad \text{be} \quad \ulcorner\langle X_1(x), X_2(x)\rangle\urcorner.$$

This is stratified with signature $\langle 2|1\rangle$ and therefore is real. Moreover, it can be proved to be a group:

(128) $$\vdash_{s} \ulcorner\mathrm{Perm}(x) \in \mathcal{V} \wedge \mathrm{Group}(\mathrm{Perm}(x))\urcorner.$$

Therefore in **ZF+NF** any real class has a permutation group which is real. In **ZF+M-Comp** we could only prove this for sets.

Finally we want to show that the category of all categories exists in **ZF+NF**.

We define a category to consist of classes $\mathcal{O}$ (objects), $\mathcal{M}$ (morphisms), and mappings Do (domain), Ra (range), $\varkappa$ (composition) which satisfy certain axioms.

For better readability we write these symbols in place of variables and give indices as superscripts which are to show the stratification.

$\ulcorner\Phi_1(\mathcal{M}, \varkappa)\urcorner$ is $\ulcorner(\mathrm{Function}(\varkappa^1) \wedge \mathrm{dom}\,\varkappa^1 \subseteq \mathcal{M}^1 \times \mathcal{M}^1 \wedge \mathrm{ran}\,\varkappa^1 \subseteq \mathcal{M}^1\urcorner$, Φ_1 is stratified of signature $\langle 1, 1\rangle$,

$\ulcorner\Phi_2(\mathcal{O}, \mathcal{M}, \mathrm{Do})\urcorner$ is $\ulcorner\mathrm{Do}^1|\mathcal{M}^1 \to \mathcal{O}^1\urcorner$,

$\ulcorner\Phi_3(\mathcal{O}, \mathcal{M}, \mathrm{Ra})\urcorner$ is $\ulcorner\mathrm{Ra}^1|\mathcal{M}^1 \to \mathcal{O}^1\urcorner$,

Φ_2, Φ_3 are stratified of signature $\langle 1, 1, 1\rangle$,

$\ulcorner\Phi_4(\mathcal{M}, \mathrm{Do}, \mathrm{Ra}, \varkappa)\urcorner$ is $\ulcorner\forall x\, \forall y\big(\langle x^0, y^0\rangle \in \mathrm{dom}\,\varkappa^1 \leftrightarrow x^0, y^0 \in \mathcal{M}^1 \wedge \mathrm{Do}^1(x^0) = \mathrm{Ra}^1(y^0)\big)\urcorner$, Φ_4 is stratified of signature $\langle 1, 1, 1, 1\rangle$,

$\ulcorner\Phi_5(\mathcal{M}, \mathrm{Do}, \mathrm{Ra}, \varkappa)\urcorner$ is $\ulcorner\forall x\, \forall y\, \forall z\big(x^0, y^0, z^0 \in \mathcal{M}^1 \wedge \mathrm{Do}^1(x^0) = \mathrm{Ra}^1(y^0) \wedge \mathrm{Do}^1(y^0) = \mathrm{Ra}^1(z^0) \to \varkappa^1\big(\varkappa^1(x^0, y^0), z^0\big) = \varkappa^1\big(x^0, \varkappa^1(y^0, z^0)\big)\big)\urcorner$, Φ_5 is stratified of singature $\langle 1, 1, 1, 1\rangle$.

$\ulcorner\Phi_6(\mathcal{O}, \mathcal{M}, \mathrm{Do}, \mathrm{Ra}, \varkappa)\urcorner$ is $\ulcorner\forall x\Big(x^0 \in \mathcal{O}^1 \to \exists y\,\big(y^0 \in \mathcal{M}^1 \wedge \mathrm{Do}^1(y^0) = x^0 \wedge \mathrm{Ra}^1(y^0) = x^0 \wedge \forall z\big(z^0 \in \mathcal{M}^1 \wedge \mathrm{Do}^1(z^0) = x^0 \to \varkappa^1(z^0, y^0) = z^0\big) \wedge \forall z\big(z^0 \in \mathcal{M}^1 \wedge x^0 = \mathrm{Ra}^1(z^0) \to \varkappa^1(y^0, z^0) = z^0\big)\big)\Big)\urcorner$, Φ_6 is stratified of signature $\langle 1, 1, 1, 1, 1\rangle$.

Let

$\ulcorner\mathrm{category}(x_1, x_2, x_3, x_4, x_5)\urcorner$ be $\ulcorner\Phi_1(x_2, x_5) \wedge \Phi_2(x_1, x_2, x_3) \wedge \Phi_3(x_1, x_2, x_4) \wedge \Phi_4(x_2, x_3, x_4, x_5) \wedge \Phi_5(x_2, x_3, x_4, x_5) \wedge \Phi_6(x_1, x_2, x_3, x_4, x_5)\urcorner$

and let

$\ulcorner\text{Category}(z)\urcorner$ be $\ulcorner\exists x_1 \ldots \exists x_5 \big(z = \langle x_1, \ldots, x_5\rangle \wedge \text{category}(x_1, \ldots, x_5)\big)\urcorner$.

Then $\ulcorner\text{category}(x_1, \ldots, x_5)\urcorner$ is stratified of signature $\langle 1, 1, 1, 1, 1\rangle$ and $\ulcorner\text{Category}(z)\urcorner$ is stratified of signature $\langle 1\rangle$.

Let CAT be $\ulcorner\{z \mid \text{Category}(z)\}\urcorner$. Then we have:

$$\vdash_3 \ulcorner\text{CAT} \in \mathfrak{V}\urcorner.$$

So the class of all categories exists in **ZF+NF.**

Next we consider functors. For better readability we write instead of variables:

$F_{\mathcal{O}}$ (mapping of objects), $F_{\mathcal{M}}$ (mapping of morphisms), cat_1, cat_2 (two categories), $\mathcal{O}_1$, $\mathcal{M}_1$, Do_1, Ra_1, $\varkappa_1$ (constituents of cat_1), $\mathcal{O}_2$, $\mathcal{M}_2$, Do_2, Ra_2, $\varkappa_2$ (constituents of cat_2).

Let

$$\ulcorner\Phi_7(\text{cat}_1, \text{cat}_2, \mathcal{O}_1, \mathcal{M}_1, \text{Do}_1, \text{Ra}_1, \varkappa_1, \mathcal{O}_2, \mathcal{M}_2, \text{Do}_2, \text{Ra}_2, \varkappa_2)\urcorner$$

be $\ulcorner\text{category}(\text{cat}_1^1) \wedge \text{category}(\text{cat}_2^1) \wedge \text{cat}_1^1 = \langle \mathcal{O}_1^1, \mathcal{M}_1^1, \text{Do}_1^1, \text{Ra}_1^1, \varkappa_1^1\rangle \wedge \text{cat}_2^1$

$$= \langle \mathcal{O}_2^1, \mathcal{O}_2^1, \text{Do}_2^1, \text{Ra}_2^1, \varkappa_2^1\rangle\urcorner.$$

This is stratified of signature $\langle 1, \ldots, 1\rangle$.

Let

$\ulcorner\Phi_8(F_{\mathcal{O}}, F_{\mathcal{M}}, \mathcal{O}_1, \mathcal{M}_1, \mathcal{O}_2, \mathcal{M}_2)\urcorner$ be $\ulcorner F_{\mathcal{O}}^1 \mid \mathcal{O}_1^1 \to \mathcal{O}_2^1 \wedge F_{\mathcal{M}}^1 \mid \mathcal{M}_1^1 \to \mathcal{M}_2^1\urcorner$.

This is stratified of signature $\langle 1, \ldots, 1\rangle$.

Let

$$\ulcorner\Phi_9(F_{\mathcal{O}}, F_{\mathcal{M}}, \mathcal{O}_1, \mathcal{M}_1, \text{Do}_1, \text{Ra}_1, \varkappa_1, \mathcal{O}_1, \mathcal{M}_2, \text{Do}_2, \text{Ra}_2, \varkappa_2)\urcorner$$

be

$$\begin{aligned} &\ulcorner\forall x \forall y \big(x^0, y^0 \in \mathcal{M}_1^1 \wedge \text{Do}_1^1(x^0) = \text{Ra}_1^1(y^0) \to F_{\mathcal{M}}^1\big(\varkappa_1^1(x^0, y^0)\big) \\ &= \varkappa_2^1\big(F_{\mathcal{M}}^1(x^0), F_{\mathcal{M}}^1(y^0)\big)\big) \wedge \forall x \big(x^0 \in \mathcal{M}_1^1 \to F_{\mathcal{O}}^1\big(\text{Do}_1^1(x^0)\big) = \text{Do}_2^1\big(F_{\mathcal{M}}^1(x^0)\big) \\ &\qquad \wedge F_{\mathcal{O}}^1\big(\text{Ra}_1^1(x^0)\big) = \text{Ra}_2^1\big(F_{\mathcal{M}}^1(x^0)\big)\big)\urcorner. \end{aligned}$$

This is stratified of signature $\langle 1, \ldots, 1\rangle$.

Finally let $\ulcorner\text{functor}(x_1, x_2, x_3, x_4)\urcorner$ be

$$\begin{aligned} &\ulcorner\exists y_1 \exists y_2 \exists y_3 \exists y_4 \exists y_5 \exists z_1 \exists z_2 \exists z_3 \exists z_4 \exists z_5 \big(\Phi_7(x_3, x_4, y_1, y_2, y_3, y_4, y_5, z_1, z_2, z_3, \\ &z_4, z_5) \wedge \Phi_8(x_1, x_2, y_1, y_2, z_1, z_2) \wedge \Phi_9(x_1, x_2, y_1, y_2, y_3, y_4, y_5, z_1, z_2, z_3, z_4, z_5)\big)\urcorner \end{aligned}$$

and let

$\ulcorner\text{Functor}(z)\urcorner$ be $\ulcorner\exists x_1 \exists x_2 \exists x_3 \exists x_4 \big(z = \langle x_1, x_2, x_3, x_4\rangle \wedge \text{functor}$
$(x_1, x_2, x_3, x_4)\big)\urcorner$.

Then $\ulcorner\text{functor}(x_1, x_2, x_3, x_4)\urcorner$ is stratified of signature $\langle 1, 1, 1, 1\rangle$ an $\ulcorner\text{Functor}(z)\urcorner$ is stratified of signature $\langle 1\rangle$.

Let FCTR be ⌜$\{z \mid \mathrm{Functor}(z)\}$⌝. Then we have:

$$\vdash_3 \ulcorner \mathrm{FCTR} \in \mathfrak{V} \urcorner .$$

This shows that the class of all functors exists in **ZF+NF**. Let DOM be ⌜$\mathrm{pr}_{4,3} \upharpoonright \mathrm{FCTR}$⌝ and RAN be ⌜$\mathrm{pr}_{4,4} \upharpoonright \mathrm{FCTR}$⌝.

Then we have:

$$\vdash_3 \ulcorner \mathrm{DOM}, \mathrm{RAN} \in \mathfrak{V} \urcorner .$$

Finally let ⌜x is composition of y, z⌝ be

$$\ulcorner \mathrm{Functor}(x) \wedge \mathrm{Functor}(y) \wedge \mathrm{Functor}(z) \wedge \mathrm{pr}_{4,3}(y) = \mathrm{pr}_{4,4}(z) \wedge \mathrm{pr}_{4,1}(x) = \mathrm{pr}_{4,1}(y) \circ \mathrm{pr}_{4,1}(z) \wedge \mathrm{pr}_{4,2}(x) = \mathrm{pr}_{4,2}(y) \circ \mathrm{pr}_{4,2}(z) \urcorner .$$

This is stratified of signature $\langle 1, 1, 1 \rangle$. If

COMP is ⌜$\{x, y, z \mid x$ is composition of $y, z\}$⌝,

then we have

$$\vdash_3 \ulcorner \mathrm{COMP} \in \mathfrak{V} \urcorner .$$

It is well known that the categories (as objects) and the functors (as morphisms) form under the natural composition a category. So we have:

$$\vdash_3 \ulcorner \mathrm{category}(\mathrm{CAT}, \mathrm{FCTR}, \mathrm{DOM}, \mathrm{RAN}, \mathrm{COMP}) \urcorner .$$

This shows that in **ZF+NF** the category of all categories exists and is itself a category:

$$\vdash_3 \ulcorner \langle \mathrm{CAT}, \mathrm{FCTR}, \mathrm{DOM}, \mathrm{RAN}, \mathrm{COMP} \rangle \in \mathrm{CAT} \urcorner . \tag{129}$$

In **ZF+M-Comp** (using ordered pairs with signature $\langle 1 \mid 1, 1 \rangle$) the formulas ⌜Category$(z)$⌝, ⌜Functor$(z)$⌝ would also be regular of signature $\langle 1 \rangle$. But putting this into the abstraction operator would require a relativization to M. Therefore in **ZF+M-Comp** only the category of all small categories that are sets exists. And this category is not small and so does not belong to itself. The theory **ZF+NF** thus proves to be a set theory fulfilling some of the demands of category theorists. But in category theory there are also some unstratified notions (e. g. Hom-functors) which can not be treated adequately by **ZF+NF**. Furthermore we do not know anything about the relative consistency of **ZF+NF** to other known systems. However, Jensen [11] has given a consistency proof for a theory **NFU** (**NF** with urelements admitted). He has pointed out that the axioms of an analogous system **ZFU+NF** (admitting urelements) are true in his model for **NFU**.

8. A generalization of stratification

A merit of **ZF+M-Comp** is the flexibility of the notion of a M-expression. Many unstratified conditions define real classes. A drawback is that the x in $\ulcorner\{x|\varphi\}\urcorner$ must always be a E-variable or relativized to M. In this respect **ZF+NF** is more favourable. But the notion of stratification is rather rigid and set formation and class formation in **ZF+NF** are quite unconnected. The question arises whether we can find a system combining the good features of both systems. For some time we believed that both systems were compatible and looked for a common extension. But we know that no consistent extension exists (cf. (103), (126)).

However, consider the following variant of **ZF+M-Comp**.

(130) *In* (78) *replace "all variables getting index* 0 *and* 1 *are to be relativized to* M*" by "all variables getting index* 0 *at some place in* φ *are to be relativized to* M*".*

This means that not only the M-variables but also the E-variables are to be relativized to M. The resulting system will be called **ZF+M-Comp-**. This is a rather weak system. We can not even prove theorems like:

$$\ulcorner x \cup y, x \cap y, -x, \bigcup x, \bigcap x, \mathcal{V} \in \mathcal{V}\urcorner.$$

Theorems like:

$$\ulcorner\emptyset, \{x, y\} \in \mathcal{V}\urcorner$$

are only derivable because of the set axioms. But (77) still holds true: any set theoretical condition defines a real class. However, note that the condition $\ulcorner x \in V\urcorner$ is not a set theoretical condition. For the formulation of the notion of transitivity one needs an unrestricted E-variable. Finally we can no longer prove $\ulcorner\langle x \mapsto \{x\}\rangle \in \mathcal{V}\urcorner$. We only get $\ulcorner\langle x \mapsto \{x\} | x \in \mathrm{M}\rangle \in \mathcal{V}\urcorner$.

But this suggests that **ZF+M-Comp-** might be compatible with **ZF+NF**. It is true that **ZF+M-Comp-** is rather uninterestingly weak. But if we combine it with **NF**, then many of the theorems of **ZF+M-Comp** which we miss in **ZF+M-Comp-** come in again by the stratification schema and a rather strong theory results. Therefore a theory **ZF+M-Comp-+NF** (if consistent) would have all the merits of **ZF+NF** and in addition get some advantages of **ZF+M-Comp** concerning unstratified conditions. Instead of adding to **ZF+Cls** two rather unrelated comprehension principles (namely **M-Comp-** and **Strat**) we now try to fuse them into one comprehension schema. For easier formulation we admit only few types of atomic formulas.

(131) *Let* φ *be a simple and usual formula containing only quasiatomic formulas of the types* $\ulcorner x \in y\urcorner$, $\ulcorner x \in \mathrm{M}\urcorner$, $\ulcorner x = y\urcorner$, $\ulcorner z = \langle x, y\rangle\urcorner$. *Suppose that to all variable occurrences in* φ *(other than immediately*

after a quantifier) is given an index which is a non-negative integer in such a way that:

(i) *the indicated occurrence of x in a subformula $\ulcorner x \in y \urcorner$ gets an index one lower than the indicated occurrence of y in that subformula;*

(ii) *the indicated occurrences of x and y in a subformula $\ulcorner x = y \urcorner$ get the same index;*

(iii) *the indicated occurrences of z, x, y in a subformula $\ulcorner z = \langle x, y \rangle \urcorner$ get the same index;*

(iv) *any variable getting an index $i \geqslant 2$ at some place in φ gets everywhere in φ the same index i in φ.*

Then $\varphi^{[\mathrm{M}]}$ comes from φ by relativizing to M *all variables getting* 0 *at some place in φ, and $\varphi^{[\mathrm{M}]}$ is called* M*-stratified.*

To give a better understanding of this notion, consider a many-sorted language with variables for hereditary sets, for classes of such sets, for classes of such classes etc. A formula of such a language may be called a n-th order set theoretical condition. A M-stratified formula corresponds to such a formula if one parallels the variables getting index 0 (i. e. index 0 only or both indices 0 and 1) to the set variables; the variables getting index 1 only to the class-of-sets-variables; the variables getting index 2 to the class-of-class-of-sets-variables etc. So a M-stratified formula may approximately be described as a formula which is "like a n-th order set theoretical condition". However, note that a M-stratified formula is a formula of a truly one-sorted language.

The M-stratification schema is:

(132) **M-Strat**: *Let φ be usual and simple and only with quasiatomic formulas of types: $\ulcorner x \in y \urcorner$, $\ulcorner x \in \mathrm{M} \urcorner$, $\ulcorner x = y \urcorner$, $\ulcorner z = \langle x, y \rangle \urcorner$. Let z be different from x and not occurring in φ. Then*

$$\ulcorner \forall \ldots \exists z \forall x (x \in z \leftrightarrow \varphi)^{[\mathrm{M}]} \urcorner$$

is an axiom if it is M*-stratified.*

ZF+M-Strat is the system which has the axioms of **ZF+Cls** and in addition all instances of **M-Strat**. $\Sigma \vdash_4 \varphi$ indicates that $\Sigma \cup A \vdash \varphi$ for some set of (closed) formulas which are axioms of **ZF+M-Strat**. Thus $\vdash_4 \varphi$ means that φ is a theorem of **ZF+M-Strat**. **M-Strat** is an extension of the stratification schema. For, if there are no indices 0, then **M-Strat** reduces to **Strat**, and by adding a constant number to all indices in a stratified formula, we can always avoid index 0.

If there are no indices greater than 1, then **M-Strat** reduces to **M-Comp**- (with pairs having also signature $\langle 1, 1, 1 \rangle$).

So **ZF+M-Strat** is a common extension of the system **ZF+NF** and **ZF+M-Comp-**. This is the most comprehensive set theory over classes suitable for the categories we know so far. But the question of relative consistency to other systems is open. Since several earlier attempts in finding a notion of M-stratification resulted in inconsistencies, we are not inclined to give a too confident prognosis. But it will have become clear that the problem of finding suitable set theories over classes is quite promising.

9. Historical remark

The desire of treating proper classes as elements has come up in mathematics in category theory. Standard examples of categories are proper classes. Standard examples of functor categories are classes of proper classes. The need for a more comprehensive set theory has been formulated several times, e. g. by McLane [13] and Ehresmann [5]. Dedecker in [4] gave some requirements that a good class theory should have, e. g. any condition on sets should determine a class (this is false in our **ZF+Cls**). Quine in his book [25] also makes some remarks towards sets of ultimate classes (see p. 321, 322 of first edition and p. 321 of second edition). An extended set theory has been given by Friedman [7] and another theory by Osius [19]. The first to bring set theory and stratification together seems to have been Houdebine [9], [10]. This system is also used by Da Costa [3]. Stratification seems also to be involved in the approach of Engeler and Röhrl [6]. Another approach is by the universes introduced by Grothendieck and Sonner (see [29]). Universes are (hereditary) sets that are models of set theory and therefore, so to speak, contain already everything of mathematical interest.

Quite another system is the system of Ackermann [1]. There are classes of non-sets in his theory, but sets are hereditary. Actually, this comes because some classes that *are* sets (in our sense) are not *called* sets.

Again of a new type are the systems **NF** and **ML** of Quine [20], [21]. However, from our intuitive picture of sets we would say that these are class theories and not set theories.

In any system discussed so far either **A 1** is false – and so there are sets of objects not being objects again – or there is no real universal class containing every object – or "set" is understood in a too narrow sense (hereditary sets) or in a too wide sense ($\mathfrak{V}$ is counted a set by Quine).

The system **ZF+Cls** and **ZF+M-Comp** had been given essentially in [14]. However, we think that our present treatment is more elegant. At the time of the writing of [14] we did not know the language with virtual terms. We now separate the material of [14] into three systems. The logical part, using ordered pairs as primitive, is given as the system **CRF** in advance. This is the underlying logic for different set theories and class theories. Such a logical system (with primitive ordered pairs) is also given by Bernays in [2] (Part II, Chapter I). The difference is that classes and coextensional objects are not identified (and so classes are always virtual in [2]) and that the language is not so flexible.

The set theoretical part is given as the system **ZF+Cls** (which is a special set theory formulated in **CRF**). This is the underlying set theory for different set theories over classes. The class theoretical part is given as the system **ZF+M-Comp** (which is a special set theory over classes).

In [14] we used M as an additional primitive. We first used ordered pairs as primitive in [17] and were forced to do so in our attempt to find the notion of M-stratification.

We are very much influenced by the writings of Quine. But we think it unfortunate to try – as he does – to reduce all of set theory, class theory, and relation theory (and even urelements) to one primitive notion ϵ.

A combination of set theory and stratification has been considered by us already in 1963 (see [16], p. 42, 43) but we did not work out the system before 1969. The work of Houdebine, Da Costa, Engeler, Röhrl has come to our attention only recently. The system of Houdebine and Da Costa is different from our system **ZF+NF** by another understanding of the notion of set. The word "set" in [9] and [3] means the same as "hereditary set" (in our sense). Accordingly there *are* real sets (in our sense) in the system of [9] and [3] which are not called "set", e. g. $\{V\}$, $\{\mathfrak{V}\}$. However, we take **ZF+Cls** as basis theory, and this guarantees that *all* sets of objects and not only some of these sets are treated by the set theoretical part of the system. Our system **ZF+NF** had been found in 1969 and for the first time been presented in Oslo in September, 1969 (together with a first version of **ZF+M-Strat**). The definition of M-stratification which we present now (and which had some inconsistent predecessors) was found recently.

References

[1] W. Ackermann, *Zur Axiomatik der Mengenlehre*, Math. Ann. 131 (1956), p. 336-345.

[2] P. Bernays and A. A. Fraenkel, *Axiomatic set theory*, Amsterdam 1958.

[3] N. Da Costa, *On a set theory suggested by Dedecker and Ehresmann*, I, II, Proc. Japan Acad. 45 (1969), p. 880-888.

[4] P. Dedecker, *Introduction aux structures locales*, Colloque de géometrie différentielle globale, Brussels 1958, p. 103-135.

[5] Ch. Ehresmann, *Gattungen von lokalen Strukturen*, Jahresbericht d. D. Math. Ver. 60 (1957), p. 49-77.

[6] E. Engeler and H. Röhrl, *On the problem of foundations of category theory*, Dialectica 23 (1969), p. 58-66.

[7] J. I. Friedman, *Proper classes as members of extended sets*, Math. Ann. 183 (1969), p. 232-240.

[8] K. Gödel, *What is Cantor's Continuum Hypothesis*, Amer. Math. Monthly 54 (1947), p. 515-525.

[9] J. Houdebine, *Classes et ensembles*, Séminaire Ch. Ehresmann, VI (1963).

[10] — *Théorie des classes et théorie des catégories*, Thèse, Rennes, 1967.

[11] R. B. Jensen, *On the consistency of a slight (?) modification of Quine's new foundations*, in: *Words and objections*. Essays on the work of W. V. Quine. Edited by D. Davidson and J. Hintikka, Dordrecht 1969.

[12] M. Kühnrich, *Zur Definition des geordneten Paares*, Z. Math. Logik Grundlagen. Math. 13 (1967), p. 379-380.

[13] S. McLane, *Locally small categories and the foundations of set theory*. In: *Infinitistic methods*, Warszawa 1961. Proceedings of a 1959 conference, p. 25-43.

[14] A. Oberschelp, *Eigentliche Klassen als Urelemente in der Mengenlehre*, Math. Ann. 157 (1964), p. 234-260.

[15] — *Sets and non-sets in set theory*, Abstract, J. Symbolic Logic 29 (1964), p. 227.

[16] — *Axiomatische Mengenlehre*, Mimeographed lecture notes, Hannover 1964.

[17] — *Mengen, Relationen, Funktionen (ein einfaches für die Schule ausreichendes System der Mengenlehre)*, Math. Phys. Semesterber. 14 (1967), p. 1-40.

[18] — *A combination of set theory and stratification*, Paper contributed to the IVth International Congress for Logic, Methodology, and Philosophy of Science, Bucharest 1971.

[19] G. Osius, *Eine Erweiterung der NBG-Mengenlehre als Grundlage der Kategorientheorie*, Mimeographed, Bielefeld 1970.

[20] W. V. Quine, *New foundations for mathematical logic*, Amer. Math. Monthly 44 (1937), p. 70-80, reprinted in [24].

[21] — *Mathematical logic*, Cambridge, Mass. 1940, rev. ed. 1951.

[22] — *On ordered pairs*, J. Symbolic Logic 10 (1945), p. 95-96.

[23] — *On what there is*, Reviews of Metaphysics (1948), reprinted in [24].

[24] — *From a logical point of view*, 9 Logico-Philosophical Essays, Cambridge, Mass. 1953.

[25] — *Set theory and its logic*, Cambridge, Mass. 1963, rev. ed. 1969.
[26] J. B. Rosser, *Logic for mathematicians*, New York, Toronto, London 1953.
[27] J. E. Rubin, *Set theory for the mathematician*, San Francisco 1967.
[28] J. Schmidt, *Mengenlehre I*, Mannheim 1966.
[29] J. Sonner, *On the formal definition of categories*, Math. Z. 80 (1962), p. 163-176.
[30] P. Suppes, *Axiomatic set theory*, Princeton 1960.
[31] H. Wang, *On Zermelo's and von Neumann's axioms for set theory*, Proc. Nat. Acad. Sci. U.S.A. 35 (1949), p. 150-155.