Formulas.

In this note we say what is meant by a formal mathematical statement. We first begin by specifying a *language* (called L), by which we mean specifying which predicate symbols ($\mathcal{P} = \{=, <, \ldots\}$), which operation symbols ($\mathcal{O} = \{+, \cdot, -, \ldots\}$), and which constant symbols ($\mathcal{C} = \{0, 1, \pi, \ldots\}$) we need for the ideas we want to express.

Example 1. (1) The language of *set theory* has one predicate symbol $\in$, no operation symbols, and no constant symbols.

(2) One language for *number theory* (i.e., the theory of the natural numbers) has one operation symbol, S (for successor), one constant symbol, 0 (for zero), and no nonlogical predicate symbols.

(3) One language for the real numbers has operation symbols $\mathcal{O} = \{+, \cdot, -\}$, constant symbols $\mathcal{C} = \{0, 1\}$, and predicate symbols $\mathcal{P} = \{<\}$.

Fixing L, we can define terms, atomic formulas, then arbitrary formulas in this language.

Definition 2. The set of all *L-terms* is the smallest set $\mathcal{T}$ such that

(i) $\mathcal{T}$ contains all variables and constant symbols, and

(ii) if $f \in \mathcal{O}$ is an n-ary operation symbol and $t_1, \ldots, t_n \in \mathcal{T}$, then $f(t_1, \ldots, t_n) \in \mathcal{T}$.

Example 3. (1) In the language of set theory the only terms are variables.

(2) In the language of number theory whose nonlogical symbols are 0 and S, the only terms are of the form $S^k(0)$ and $S^k(x_i)$, $k = 0, 1, 2, \ldots$.

(3) In the language of the real numbers whose nonlogical symbols are $+, \cdot, -, 0, 1, <$ there are very complicated terms like $(((x_1 \cdot x_{17}) + ((x_1 \cdot 0) \cdot x_9)) + 1)$.

Definition 4. The set of all *atomic L-formulas* is the set of all strings $P(t_1, \ldots, t_n)$ where P is an n variable predicate symbol and the t_i are terms.

Example 5. (1) In the language of set theory the only atomic formulas are of the form $(x_i \in x_j)$.

(2) In the language of number theory whose nonlogical symbols are 0 and S, the only atomic formulas are equations of the form $(S^k(x_i) = S^\ell(x_j))$, $(S^k(x_i) = S^\ell(0))$, $(S^k(0) = S^\ell(x_j))$, and $(S^k(0) = S^\ell(0))$.

(3) In the language of the real numbers whose nonlogical symbols are $+, \cdot, -, 0, 1, <$ there are very complicated atomic formulas, including $(1 < (x \cdot x))$ or $((x_1 + (x_2 + x_3)) = ((x_1 + x_2) + x_3))$.

Definition 6. The set of all *L-formulas* is the smallest set $\mathcal{F}$ such that

(i) $\mathcal{F}$ contains all atomic formulas, and

(ii) if $\alpha, \beta \in \mathcal{F}$ and x is a variable, then the following are in $\mathcal{F}$: $(\alpha \wedge \beta)$, $(\alpha \vee \beta)$, $(\alpha \rightarrow \beta)$, $(\alpha \leftrightarrow \beta)$, $(\neg\alpha)$, $(\forall x\ \alpha)$, $(\exists x\ \alpha)$.

Example 7. In any language, the formulas get complicated. Here are some examples.

(1) (Set theory) We can express "x is a subset of y" with the formula $\alpha(x, y) =$ "$\forall z\ ((z \in x) \to (z \in y))$".

(2) (Number theory) We can express that the successor function is 1-1 with the formula $\beta = \forall x\ \forall y\ ((S(x) = S(y)) \to (x = y))$.

(3) (Real numbers) We can express that any monic cubic polynomial has a root with the formula $\gamma = \forall y_1\ \forall y_2\ \forall y_3\ \exists x\ (x^3 + y_1 \cdot x^2 + y_2 \cdot x + y_3 = 0)$.

Exercises. Express the given fact or relation in the language whose nonlogical symbols are those given.

(1) Express "There is a set with no elements" in the language of set theory.

$$\exists x\ \forall y\ (\neg(y \in x))$$

(2) Express "x has exactly two elements" in the language of set theory.

$$\exists y\ \exists z\ (\underbrace{(y \in x) \wedge (z \in x) \wedge (\neg(y = z))}_{x \text{ has 2 distinct elements}} \wedge \underbrace{\forall w\ ((w \in x) \to ((w = y) \vee (w = z)))}_{x \text{ has no other elements}})$$

(3) One language for ordered sets has $\leq$ as its only nonlogical symbol. In this language express "x is not the largest element and not the smallest element."

$$\exists y\ \exists z\ (\quad \underbrace{(\neg(x \leq y))}_{x \text{ is not smallest}} \quad \wedge \quad \underbrace{(\neg(z \leq x))}_{x \text{ is not largest}} \quad)$$

(4) Let L be a language whose only nonlogical symbol is a binary predicate symbol E. Express the fact that "E defines an equivalence relation". (That is, express the fact that the set of all pairs (x, y) such that $E(x, y)$ is true is an equivalence relation.)

$$\underbrace{(\forall x\ E(x, x))}_{E \text{ reflexive}} \wedge \underbrace{(\forall x\ \forall y\ (E(x, y) \to E(y, x)))}_{E \text{ symmetric}} \wedge \underbrace{(\forall x\ \forall y\ \forall z\ ((E(x, y) \wedge E(y, z)) \to E(x, z)))}_{E \text{ transitive}}$$

(5) Express Fermat's Last Theorem in a language for number theory whose nonlogical symbols are $0, +, \cdot, \hat{}\,, <$. (Fermat's Last Theorem is the statement that if x, y, z, n are nonzero natural numbers and n is at least 3, then $x\hat{}n + y\hat{}n = z\hat{}n$ does not hold.)

We need to define formulas in this language that allow us to refer to natural numbers other than 0. Let $1(x) =$ "$\forall y\ (x \cdot y = y)$." When applied to a natural number x the formula $1(x)$ is true only when $x = 1$. Now define $2(x) = \exists y\ (1(y) \wedge (x = y + y))$, $3(x) = \exists y\ \exists z\ (1(y) \wedge 2(z) \wedge (x = y + z))$, etc. Now, the formula we seek is

$$\forall x\ \forall y\ \forall z\ \forall n\ ((x\hat{}n + y\hat{}n = z\hat{}n) \to ((x = 0) \vee (y = 0) \vee (z = 0) \vee (\exists w\ ((3(w) \wedge (n < w))))))$$

(6) Express "$\gcd(x, y) = 1$" in a language for number theory whose nonlogical symbols are $+, \cdot, 0, <$.

Let $1(x)$ be the formula from Exercise (5) and let $x \mid y$ be the formula "$\exists z\ (x \cdot z = y)$". With these abbreviations, a formula that works is

$$\forall z\ (((z \mid x) \wedge (z \mid y)) \to 1(z))$$

If you prefer to avoid abbreviations, the full formula is

$$\forall z\ (((\exists u\ (z \cdot u = x)) \wedge (\exists v\ (z \cdot v = y))) \to (\forall w\ (z \cdot w = w)))$$