Notre Dame Journal of Formal Logic
Volume X, Number 1, January 1969

A FORMALISATION OF THE ARITHMETIC OF THE ORDINALS LESS THAN ω^ω

H. P. WILLIAMS

Some of the results of ordinal arithmetic can be derived from a multi-successor equation calculus. The initial functions are:

(i) the zero function $\mathsf{N}(x) = 0$
(ii) the identity function $\mathsf{I}(x) = x$.

These two functions are implicit. In addition there are:

(iii) a countable number of successor functions $S_0, S_1, S_2, \ldots$.

The successor functions are restricted by the axioms

A $$S_\mu S_\nu = S_\mu \text{ if } \mu > \nu$$
B $$S_a S_b \ldots S_q = S_{a'} S_{b'} \ldots S_{q'}$$

with $a \leq b \leq \ldots \leq q$ and $a' \leq b' \leq \ldots \leq q'$ if and only if $a = a'$, $b = b'$, $\ldots q = q'$.

A function may be defined explicitly, or by recursion in the following way

$$F(x, 0) = a(x)$$
$$F(x, S_\mu y) = b_\mu(x, y, F(x, y))$$

from previously defined functions $a(x)$ and $b_\mu(x, y, z)$ (for all μ) if the b_μ obey the following identity imposed by **A**:

C $$b_\mu(x, S_\nu y, b_\nu(x, y, z)) = b_\mu(x, y, z) \text{ if } \nu < \mu,$$

The rules of inference are the following schemata

Sb$_1$ $$\frac{F(x) = G(x)}{F(A) = G(A)}$$

Sb$_2$ $$\frac{A = B}{F(A) = F(B)}$$

Received October 30, 1967

T
$$\frac{A = B \quad A = C}{B = C}$$

and the uniqueness rule

U
$$\frac{F(\mathsf{S}_\mu x) = \mathsf{H}_\mu(x, F(x))}{F(x) = \mathsf{H}^x F(0)} \text{ for all } \mu$$

F, G, H_μ are recursive functions and A, B, C are recursive terms. $\mathsf{H}^x t$ is defined by the primitive recursion $\mathsf{H}^0 t = t$, $\mathsf{H}^{c_\mu x} t = \mathsf{H}_\mu(x, \mathsf{H}^x t)$. **U** may be shown to be equivalent to the schema

U₁
$$\frac{f(0) = g(0) \quad f(\mathsf{S}_\mu x) = \mathsf{H}_\mu(x, f(x)) \text{ for all } \mu \quad g(\mathsf{S}_\mu x) = \mathsf{H}_\mu(x, g(x))}{f(x) = g(x)}$$

$\mathsf{S}_\mu 0$ is interpreted as $\omega^\mu \cdot \omega^0$ is understood to be 1 and S_0 generates the natural numbers starting with 0. Addition is defined by the following recursion:

$$a + 0 = a,\ a + \mathsf{S}_\mu b = \mathsf{S}_\mu(a + b)\ .$$

Predecessor functions P_0, P_1, P_2, . . . are introduced by the following definitions:

(i) $\mathsf{P}_\mu 0 = 0$ for all μ
(ii) $\mathsf{P}_\mu \mathsf{S}_\nu a = \mathsf{P}_\mu a$ if $\nu < \mu$
(iii) $\mathsf{P}_\mu \mathsf{S}_\nu a = \mathsf{S}_\nu a$ if $\nu > \mu$.

$\mathsf{P}_\mu \mathsf{S}_\mu a$ is defined by the following

(iv) $\mathsf{P}_\mu \mathsf{S}_\mu 0 = 0$
(v) $\mathsf{P}_\mu \mathsf{S}_\mu \mathsf{S}_\nu a = \mathsf{P}_\mu \mathsf{S}_\mu a$ if $\nu < \mu$
(vi) $\mathsf{P}_\mu \mathsf{S}_\mu \mathsf{S}_\mu a = \mathsf{S}_\nu a$ if $\nu \geq \mu$

We must verify that these definitions obey the consistency condition **C**. Consider $\mathsf{P}_\mu \mathsf{S}_\nu \mathsf{S}_\lambda a$ when $\nu > \lambda$

Case (1) $\mu < \nu$

$$\begin{aligned} \mathsf{P}_\mu \mathsf{S}_\nu \mathsf{S}_\lambda a &= \mathsf{S}_\nu \mathsf{S}_\lambda a \text{ by (ii)} \\ &= \mathsf{S}_\nu a \\ \mathsf{P}_\mu \mathsf{S}_\nu a &= \mathsf{S}_\nu a \text{ by (ii)} \end{aligned}$$

Case (2) $\mu = \nu$

$$\mathsf{P}_\mu \mathsf{S}_\nu \mathsf{S}_\lambda a = \mathsf{P}_\mu \mathsf{S}_\mu \mathsf{S}_\lambda a = \mathsf{P}_\mu \mathsf{S}_\mu a \text{ by (v)}$$
$$\mathsf{P}_\mu \mathsf{S}_\nu a = \mathsf{P}_\mu \mathsf{S}_\mu a$$

Case (3) $\mu > \nu$

$$\mathsf{P}_\mu \mathsf{S}_\nu \mathsf{S}_\lambda a = \mathsf{P}_\mu \mathsf{S}_\lambda a = \mathsf{P}_\mu a \text{ by (iii)}$$
$$\mathsf{P}_\mu \mathsf{S}_\nu a = \mathsf{P}_\mu a \text{ by (iii)}$$

Subtraction is defined by the following recursions:

$$a \dot{-} 0 = a, \quad a \dot{-} \mathrm{S}_\mu b = (a \dot{-} b) \dot{-} \omega^\mu, \quad a \dot{-} \omega^\mu = \mathrm{P}_\mu a$$

It must be verified that the functions in terms of which addition and subtraction are defined obey the consistency condition **C**.

For addition

$$\mathrm{S}_\mu \mathrm{S}_\nu (a + b) = \mathrm{S}_\mu (a + b) \text{ if } \nu < \mu.$$

For subtraction it is first necessary to prove the following result.

(1a) $\mathrm{P}_\mu \mathrm{P}_\nu a = \mathrm{P}_\mu a, \ \nu < \mu$

Let $f(a) = \mathrm{P}_\mu \mathrm{P}_\nu a$, $g(a) = \mathrm{P}_\mu a$

$f(0) = \mathrm{P}_\mu \mathrm{P}_\nu 0 = \mathrm{P}_\mu 0 = g(0)$ by (i)

If $\lambda < \nu$, $f(\mathrm{S}_\lambda a) = \mathrm{P}_\mu \mathrm{P}_\nu \mathrm{S}_\lambda a = \mathrm{P}_\mu \mathrm{P}_\nu a = f(a)$ by (iii)

$g(\mathrm{S}_\lambda a) = \mathrm{P}_\mu \mathrm{S}_\lambda a = \mathrm{P}_\mu a = f(a)$ by (iii)

If $\lambda > \nu$, $f(\mathrm{S}_\lambda a) = \mathrm{P}_\mu \mathrm{P}_\nu \mathrm{S}_\lambda a = \mathrm{P}_\mu \mathrm{S}_\lambda = g(\mathrm{S}_\lambda a)$ by (ii)

If $\lambda = \nu$, $f(\mathrm{S}_\lambda a) = \mathrm{P}_\mu \mathrm{P}_\nu \mathrm{S}_\nu a$

$g(\mathrm{S}_\lambda a) = \mathrm{P}_\mu \mathrm{S}_\nu a$

Let $p(a) = f(\mathrm{S}_\lambda a)$, $q(a) = g(\mathrm{S}_\lambda a)$.

Then $p(0) = \mathrm{P}_\mu \mathrm{P}_\nu \mathrm{S}_\nu 0 = \mathrm{P}_\mu 0 = 0$

$q(0) = \mathrm{P}_\mu \mathrm{S}_\nu 0 = 0$

If $k \geq \nu$ $p(\mathrm{S}_k a) = \mathrm{P}_\mu \mathrm{P}_\nu \mathrm{S}_\nu \mathrm{S}_k a = \mathrm{P}_\mu \mathrm{S}_k a$ by (vi)

$q(\mathrm{S}_k a) = \mathrm{P}_\mu \mathrm{S}_\nu \mathrm{S}_k a = \mathrm{P}_\mu \mathrm{S}_k a$ by (iii)

If $k < \nu$ $p(\mathrm{S}_k a) = \mathrm{P}_\mu \mathrm{P}_\nu \mathrm{S}_k a = \mathrm{P}_\mu \mathrm{S}_\nu \mathrm{S}_\nu a = \mathrm{P}(a)$ by (v)

$q(\mathrm{S}_k a) = \mathrm{P}_\mu \mathrm{S}_\nu \mathrm{S}_k a = \mathrm{P}_\mu \mathrm{S}_k a$ by (iii)

$= \mathrm{P}_\mu a$ by (iii)

$= \mathrm{P}_\mu \mathrm{S}_\nu a$ by (iii)

$= q(a)$

We will now prove the following

(1b) $\mathrm{P}_\mu \mathrm{P}_\nu a = \mathrm{P}_\mu a, \ \nu < \mu$

Let $f(a) = \mathrm{P}_\nu \mathrm{P}_\mu a$, $g(a) = \mathrm{P}_\mu a$

$f(0) = \mathrm{P}_\nu \mathrm{P}_\mu 0 = \mathrm{P}_\nu 0 = 0$

$g(0) = \mathrm{P}_\mu 0 = 0$

Case (1) $\lambda < \nu < \mu$

$f(\mathrm{S}_\lambda a) = \mathrm{P}_\nu \mathrm{P}_\mu \mathrm{S}_\lambda a = \mathrm{P}_\nu \mathrm{P}_\mu a = f(a)$ by (ii)

$g(\mathrm{S}_\lambda a) = \mathrm{P}_\mu \mathrm{S}_\lambda a = \mathrm{P}_\mu a = g(a)$ by (ii)

Case (2) $\nu \leq \lambda < \mu$

$f(\mathrm{S}_\lambda a) = \mathrm{P}_\nu \mathrm{P}_\mu \mathrm{S}_\lambda a = \mathrm{P}_\nu \mathrm{P}_\mu a = f(a)$ by (ii)

$g(\mathrm{S}_\lambda a) = \mathrm{P}_\mu \mathrm{S}_\lambda a = \mathrm{P}_\mu a = g(a)$ by (ii)

Case (3) $\nu < \mu < \lambda$

$$f(S_\lambda a) = P_\nu P_\mu S_\lambda a = P_\nu S_\lambda a = S_\lambda a \text{ by (iii)}$$
$$g(S_\lambda a) = P_\mu S_\lambda a = S_\lambda a \text{ by (iii)}$$

Case (4) $\nu < \mu = \lambda$

$$f(S_\lambda a) = P_\nu P_\mu S_\mu a = m(a)$$
$$g(S_\lambda a) = P_\mu S_\mu a = n(a).$$

If $\delta < \mu$, $m(S_\delta a) = P_\nu P_\mu S_\mu S_\delta a = P_\nu P_\mu S_\mu a = m(a)$ by (v)
$n(S_\delta a) = P_\mu S_\mu S_\delta a = P_\mu S_\mu a = n(a)$ by (v)

If $\delta \geq \mu$, $m(S_\delta a) = P_\nu P_\mu S_\mu S_\delta a = P_\nu S_\delta a = S_\delta a$ by (vi) and (iii)
$n(S_\delta a) = P_\mu S_\mu S_\delta a = S_\delta a$ by (vi)

We can combine (1a) and (1b) to give

(1) $$P_\mu P_\nu a = P_\nu P_\mu a$$

The consistency of the defining equations

$$a \dot{-} S_\mu b = (a \dot{-} b) \dot{-} \omega^\mu$$

can now be proved since

$$\begin{aligned} a \dot{-} S_\mu S_\nu b &= (a \dot{-} S_\nu b) \dot{-} \omega^\mu = ((a \dot{-} b) \dot{-} \omega^\nu) \dot{-} \omega^\mu \\ &= P_\mu P_\nu (a \dot{-} b) = P_\mu (a \dot{-} b) \text{ if } \nu < \mu \\ &= (a \dot{-} b) \dot{-} \omega^\mu \\ &= a \dot{-} S_\mu b \end{aligned}$$

The degree function d. The function $\mathrm{Max}(x, y)$ on the natural numbers is taken as defined. Then the degree function d defined on the ordinals but having values only among the natural numbers is defined by the following recursion.

$$\mathrm{d}(0) = 0$$
$$\mathrm{d}(S_\mu a) = \mathrm{Max}(\mathrm{d}(a), \mu).$$

The consistency condition is satisfied since

$$\mathrm{Max}(\mathrm{Max}(\mathrm{d}(a), \nu), \mu) = \mathrm{Max}(\mathrm{d}(a), \mu) \text{ if } \nu < \mu.$$

Multiplication is defined by the following recursions.

$$\begin{aligned} a \cdot 0 &= 0 \\ a \cdot S_0 b &= a \cdot b + a \\ a \cdot S_\mu b &= a \cdot b + a \cdot \omega^\mu \quad \mu > 0 \\ 0 \cdot \omega^\mu &= 0 \\ S_\nu a \cdot \omega^\mu &= \omega^{\max(\mathrm{d}(a), \nu) + \mu} \end{aligned}$$

The consistency of the defining equations for $a \cdot \omega^\mu$ follows from the identity

$$\mathrm{Max}(\mathrm{Max}(\mathrm{d}(a), \lambda), \nu) = \mathrm{Max}(\mathrm{d}(a), \nu) \text{ if } \lambda < \nu.$$

To prove the consistency of the defining equations for $a \cdot b$ it is first necessary to prove the following results.

(2) $\omega^{\nu} + \omega^{\mu} = \omega^{\mu}$ if $\nu < \mu$

$\omega^{\nu} + \omega^{\mu} = \mathsf{S}_{\nu}0 + \mathsf{S}_{\mu}0 = \mathsf{S}_{\mu}(\mathsf{S}_{\nu}0 + 0)$

$= \mathsf{S}_{\mu}\mathsf{S}_{\nu}0 = \mathsf{S}_{\mu}0 = \omega^{\mu}$

(3) $a \cdot \omega^{\nu} + a \cdot \omega^{\mu} = a \cdot \omega^{\mu}$ if $\nu < \mu$

$0 \cdot \omega^{\nu} + 0 \cdot \omega^{\mu} = 0$

$0 \cdot \omega^{\mu} = 0$

$$\begin{aligned} \mathsf{S}_{\lambda}a \cdot \omega^{\nu} + \mathsf{S}_{\lambda}a \cdot \omega^{\mu} &= \omega^{\mathrm{Max}(\mathrm{d}(a),\lambda)+\nu} + \omega^{\mathrm{Max}(\mathrm{d}(a),\lambda)+\mu} \\ &= \omega^{\mathrm{Max}(\mathrm{d}(a),\lambda)+\mu} \text{ by (2) if } \nu < \mu \\ &= \mathsf{S}_{\lambda}a \cdot \omega^{\mu} \end{aligned}$$

The consistency can now be proved for

$$\begin{aligned} a \cdot \mathsf{S}_{\nu}b + a \cdot \omega^{\mu} &= a \cdot b + a \cdot \omega^{\mu} + a \cdot \omega^{\mu} \\ &= a \cdot b + a \cdot \omega^{\mu} \text{ if } \nu < \mu \end{aligned}$$

Some results concerning the function d are now proved.

(4) $\mathrm{d}(\omega^{\nu}) = \nu$

$\mathrm{d}(\mathsf{S}_{\nu}0) = \mathrm{Max}(\mathrm{d}(0),\nu) = \nu.$

(5) $\mathrm{d}(a + b) = \mathrm{Max}(\mathrm{d}(a),\mathrm{d}(b))$

$\mathrm{d}(a + 0) = \mathrm{d}(a)$

$\mathrm{Max}(\mathrm{d}(a),\mathrm{d}(0)) = \mathrm{Max}(\mathrm{d}(a),0) = \mathrm{d}(a)$

$\mathrm{d}(a + \mathsf{S}_{\mu}b) = \mathrm{d}(\mathsf{S}_{\mu}(a + b)) = \mathrm{Max}(\mathrm{d}(a + b),\mu)$

$$\begin{aligned} \mathrm{Max}(\mathrm{d}(a),\mathrm{d}(\mathsf{S}_{\mu}b)) &= \mathrm{Max}(\mathrm{d}(a),\mathrm{Max}(\mathrm{d}(b),\mu)) \\ &= \mathrm{Max}(\mathrm{Max}(\mathrm{d}(a),\mathrm{d}(b)),\mu) \end{aligned}$$

The result follows by $\mathbf{U}_2$.

(6) $\mathrm{d}(\mathsf{S}_{\mu}a \cdot \mathsf{S}_{\nu}b) = \mathrm{d}(\mathsf{S}_{\mu}a) + \mathrm{d}(\mathsf{S}_{\nu}b)$

$$\begin{aligned} \mathrm{d}(\mathsf{S}_{\mu}a \cdot \mathsf{S}_{\nu}0) = \mathrm{d}(\mathsf{S}_{\mu}a \cdot \omega^{\nu}) &= \mathrm{d}(\omega^{\mathrm{d}(\mathsf{S}_{\mu}a)+\nu}) \\ &= \mathrm{d}(\mathsf{S}_{\mu}a) + \nu \text{ by (4)} \\ &= \mathrm{d}(\mathsf{S}_{\mu}a) + \mathrm{d}(\mathsf{S}_{\nu}0) \end{aligned}$$

$$\begin{aligned} \mathrm{d}(\mathsf{S}_{\mu}a \cdot \mathsf{S}_{\nu}\mathsf{S}_{\lambda}b) &= \mathrm{d}(\mathsf{S}_{\mu}a \cdot \mathsf{S}_{\lambda}b + \mathsf{S}_{\mu}a \cdot \omega^{\nu}) \\ &= \mathrm{Max}(\mathrm{d}(\mathsf{S}_{\mu}a \cdot \mathsf{S}_{\lambda}b), \mathrm{d}(\mathsf{S}_{\mu}a \cdot \omega^{\nu})) \\ &= \mathrm{Max}(\mathrm{d}(\mathsf{S}_{\mu}a \cdot \mathsf{S}_{\lambda}b),\mathrm{d}(\mathsf{S}_{\mu}a) + \nu) \\ \mathrm{d}(\mathsf{S}_{\mu}a) + \mathrm{d}(\mathsf{S}_{\nu}\mathsf{S}_{\lambda}b) &= \mathrm{d}(\mathsf{S}_{\mu}a) + \mathrm{Max}(\mathrm{d}(\mathsf{S}_{\lambda}a),\nu) \\ &= \mathrm{Max}(\mathrm{d}(\mathsf{S}_{\mu}a) + \mathrm{d}(\mathsf{S}_{\lambda}a), \mathrm{d}(\mathsf{S}_{\mu}a) + \nu) \end{aligned}$$

The result follows by $\mathbf{U}_1$.

Some results of elementary ordinal arithmetic are now proved.

Associativity of addition

(7) $(a + b) + c = a + (b + c)$

$(a + b) + 0 = a + b$

$a + (b + 0) = a + b$

$$(a + b) + S_\mu c = S_\mu((a + b) + c)$$
$$a + (b + S_\mu c) = a + S_\mu(b + c) = S_\mu(a + (b + c))$$

The result follows by $\mathbf{U}_1$.

The left distributive law

(8)
$$a \cdot (b + c) = a \cdot b + a \cdot c$$
$$a \cdot (b + 0) = a \cdot b$$
$$a \cdot b + a \cdot 0 = a \cdot b + 0 = a \cdot b$$
$$a \,.\, (b + S_\mu c) = a \cdot S_\mu(b + c) = a \cdot (b + c) + a \cdot \omega^\mu$$
$$a \cdot b + a \cdot S_\mu c = a \cdot b + (a \cdot c + a \,.\, \omega^\mu)$$
$$= (a \cdot b + a \cdot c) + a \cdot \omega^\mu \text{ by (7)}$$

The result follows by $\mathbf{U}_1$.

Before proving the associativity of multiplication the following less general result is proved.

(9)
$$a \cdot (b \cdot \omega^\mu) = (a \cdot b) \cdot \omega^\mu$$
$$a \cdot (0 \cdot \omega^\mu(= a \cdot 0 = 0$$
$$(a \cdot 0) \cdot \omega^\mu = 0 \cdot \omega^\mu = 0$$
$$a \cdot (S_\nu b \cdot \omega^\mu) = a \cdot \omega^{\mathrm{Max}(d(b),\nu)+\mu}$$
$$= a \cdot \omega^{d(S_\nu b)+\mu}$$
$$(a \cdot S_\nu b) \cdot \omega^\mu = (a \cdot b + a \cdot \omega^\nu) \cdot \omega^\mu$$

It is necessary to prove

$$a \cdot \omega^{d(S_\nu b)+\mu} = (a \cdot b + a \cdot \omega^\nu) \cdot \omega^\mu$$
$$0 \cdot \omega^{d(S_\nu b)+\mu} = 0$$
$$(0 \cdot b + 0 \cdot \omega^\nu) \cdot \omega^\mu = 0$$
$$S_\lambda a \cdot \omega^{d(S_\nu b)+\mu} = \omega^{d(S_\lambda a)+d(S_\nu b)+\mu}$$
$$(S_\lambda a \cdot b + S_\lambda a \cdot \omega^\nu) \cdot \omega^\mu = (S_\lambda a \cdot b + \omega^{d(S_\lambda a)+\nu}) \cdot \omega^\mu$$
$$= S_{d(S_\lambda a)+\nu}(S_\lambda a \cdot b) \cdot \omega^\mu$$
$$= \omega^{\mathrm{Max}(d(S_\lambda a,b),d(S_\lambda a)+\nu)+\mu}$$

It remains to show

$$d(S_\lambda a) + d(S_\nu b) + \mu = \mathrm{Max}(d(S_\lambda a \cdot b), d(S_\lambda a) + \nu) + \mu$$
$$d(S_\lambda a) + d(S_\nu 0) = d(S_\lambda a) + \nu$$
$$\mathrm{Max}(d(S_\lambda a \cdot 0), d(S_\lambda a) + \nu) = \mathrm{Max}(0, d(S_\lambda a) + \nu)$$
$$= d(S_\lambda a) + \nu$$
$$\mathrm{Max}(d(S_\lambda a \cdot S_\delta b), d(S_\lambda a) + \nu)$$
$$= \mathrm{Max}(d(S_\lambda a) + d(S_\delta b), d(S_\lambda a) + \nu) \text{ by (6)}$$
$$= d(S_\lambda a) + \mathrm{Max}(d(S_\delta b), \nu)$$
$$= d(S_\lambda a) + d(S_\nu S_\delta b)$$

Hence the result.

Associativity of Multiplication

(10)
$$a \cdot (b \cdot c) = (a \cdot b) \cdot c$$

$$a \cdot (b \cdot 0) = a \cdot 0 = 0$$
$$(a \cdot b) \cdot 0 = 0$$
$$a \cdot (b \cdot \mathrm{S}_\mu c) = a \cdot (b \cdot c + b \cdot \omega^\mu)$$
$$= a \cdot (b \cdot c) + a \cdot (b \cdot \omega^\mu)$$
$$(a \cdot b) \cdot \mathrm{S}_\mu c = (a \cdot b) \cdot c + (a \cdot b) \cdot \omega^\mu$$
$$= (a \cdot b) \cdot c + a \cdot (b \cdot \omega^\mu) \text{ by (9)}$$

The result follows by $\mathbf{U}_1$.

(11) $$0 + a = a$$
$$0 + 0 = 0$$
$$0 + \mathrm{S}_\mu a = \mathrm{S}_\mu(0 + a)$$

Component Functions These are defined by the following equations

$$\mathrm{C}_\mu(0) = 0$$
$$\mathrm{C}_\mu(\mathrm{S}_\nu a) = \mathrm{C}_\mu(a) \text{ if } \nu < \mu$$
$$\mathrm{C}_\mu(\mathrm{S}_\mu a) = \mathrm{S}_0\mathrm{C}_\mu(a)$$
$$\mathrm{C}_\mu(\mathrm{S}_\nu a) = 0 \text{ if } \nu > \mu .$$

These definitions obey the consistency condition C since

(12) $$\mathrm{C}_\mu(\mathrm{S}_\nu \mathrm{S}_\lambda a) = \mathrm{C}_\mu(\mathrm{S}_\lambda a)$$
$$= \mathrm{C}_\mu(a)$$
$$= \mathrm{C}_\mu(\mathrm{S}_\nu a) \text{ if } \lambda < \nu < \mu$$
$$\mathrm{C}_\mu(\mathrm{S}_\nu \mathrm{S}_\lambda a) = 0$$
$$= \mathrm{C}_\mu(\mathrm{S}_\nu a) \text{ if } \nu > \mu, \lambda < \nu$$
$$\mathrm{C}_\mu(\mathrm{S}_\mu \mathrm{S}_\nu a) = \mathrm{S}_0\mathrm{C}_\mu(\mathrm{S}_\nu a)$$
$$= \mathrm{S}_0\mathrm{C}_\mu(a)$$
$$= \mathrm{C}_\mu(\mathrm{S}_\mu a) \text{ if } \nu < \mu .$$

Before Cantor's Normal Form theorem is proved a number of results are required.

(13) $$\omega^\nu \cdot \mathrm{C}_\nu(a) + \omega^\mu = \omega^\mu \text{ if } \nu < \mu$$
$$\omega^\nu \cdot \mathrm{C}_\nu(0) + \omega^\mu = \omega^\nu \cdot 0 + \omega^\mu$$
$$= \omega^\mu \text{ by (11)}$$
$$\omega^\nu \cdot \mathrm{C}_\nu(\mathrm{S}_\lambda a) + \omega^\mu = \omega^\nu \cdot \mathrm{C}_\nu(a) + \mu \text{ if } \lambda < \nu$$
$$\omega^\nu \cdot \mathrm{C}_\nu(\mathrm{S}_\nu a) + \omega^\mu = \omega^\nu \cdot \mathrm{S}_0\mathrm{C}_\nu(a) + \omega^\mu$$
$$= (\omega^\nu \cdot \mathrm{C}_\nu(a) + \omega^\nu) + \omega^\mu$$
$$= \omega^\nu \cdot \mathrm{C}_\nu(a) + (\omega^\nu + \omega^\mu) \text{ by (7)}$$
$$= \omega^\nu \cdot \mathrm{C}_\nu(a) + \omega^\mu \text{ by (2)}$$
$$\omega^\nu \cdot \mathrm{C}_\nu(\mathrm{S}_\lambda a) + \omega^\mu = \omega^\nu \cdot 0 + \omega^\mu$$
$$= \omega^\mu \text{ if } \lambda > \nu \text{ by (1).}$$

The result follows by $\mathbf{U}_1$.

(14) $$a \cdot + \omega^\lambda = \omega^\lambda \cdot \mathrm{C}_\lambda(a) + \omega^\lambda \text{ if } \lambda \geq \mathrm{d}(a)$$
$$0 + \omega^\lambda = \omega^\lambda$$
$$\omega^\lambda \cdot \mathrm{C}_\lambda(0) + \omega^\lambda = \omega^\lambda \cdot 0 + \omega^\lambda$$
$$= \omega^\lambda$$

$$\begin{aligned}
S_\mu a + \omega^\lambda &= S_\lambda S_\mu a \\
&= S_\lambda a \text{ if } \mu < \lambda \\
&= a + \omega^\lambda \\
\omega^\lambda \cdot C_\lambda(S_\mu a) + \omega^\lambda &= \omega^\lambda C_\lambda(a) + \omega^\lambda \text{ if } \mu < \lambda \\
S_\lambda a + \omega^\lambda &= (a + \omega^\lambda) + \omega^\lambda \\
\omega^\lambda \cdot C_\lambda(S_\lambda a) + \omega^\lambda . &= \omega^\lambda \cdot S_0 C_\lambda(a) + \omega^\lambda \\
&= (\omega^\lambda \cdot C_\lambda(a) + \omega^\lambda) + \omega^\lambda \text{ if } \mu > \lambda \text{ and } d(S_\mu a) > \lambda.
\end{aligned}$$

The result follows by $\mathbf{U}_1$.

The Sum Function Given any recursive function $f(x)$ the function $\sum_0^n f(x)$ is defined on the natural numbers by the following recursions

$$\sum_0^0 f(x) = f(0)$$
$$\sum_0^{S_0 n} f(x) = f(S_0 n) + \sum_0^n f(x)$$

Cantor's Normal Form Theorem

(15) $$a = \sum_0^{d(a)} \omega^x \cdot C_x(a)$$

Let the right hand side be $f(a)$

$$f(0) = 0$$
$$f(S_\lambda a) = \sum_0^{d(S_\lambda a)} \omega^x \cdot C_x(S_\lambda a)$$

Case (i) $\lambda \geq d(a) \cdot d(S_\lambda a) = \text{Max}(d(a), \lambda) = \lambda$.

$$\begin{aligned}
\text{Hence } f(S_\lambda a) &= \sum_0^\lambda \omega^x \cdot C_x(S_\lambda a) \\
&= \omega^\lambda \cdot C_\lambda(S_\lambda a) + \sum_0^{\lambda-1} \omega^x \cdot C_x(S_\lambda a) \\
&= \omega^\lambda \cdot S_0 C_\lambda(a) + 0 \\
&= \omega^\lambda \cdot C_\lambda(a) + \omega^\lambda. \\
&= a + \omega^\lambda \\
&= S_\lambda a \text{ by (14)}
\end{aligned}$$

Case (ii) $\lambda < d(a)$, $d(S_\lambda a) = \text{Max}(d(a), \lambda) = d(a)$

$$\begin{aligned}
f(S_\lambda a) &= \sum_0^{d(a)} \omega^x \cdot C_x(S_\lambda a) \\
&= \sum_0^{d(a)-1-\lambda} \omega^{\lambda+1+x} \cdot C_{\lambda+1+x}(S_\lambda a) + \sum_0^\lambda \omega^x \cdot C_x(S_\lambda a) \\
&= \sum_0^{d(a)-1-\lambda} \omega^{\lambda+1+x} \cdot C_{\lambda+1+x}(a) + \omega^\lambda \cdot C_\lambda(S_\lambda a) + \sum_0^{\lambda-1} \omega^x \cdot C_x(S_\lambda a) \\
&= \sum_0^{d(a)-1-\lambda} \omega^{\lambda+1+x} \cdot C_{\lambda+1+x}(a) + \omega^\lambda \cdot S_0 C_\lambda(a)
\end{aligned}$$

$$= \sum_{0}^{d(a)-1-\lambda} \omega^{\lambda+1+x} \cdot C_{\lambda+1+x}(a) + \omega^{\lambda} \cdot C_{\lambda}(a) + \omega^{\lambda}$$
$$= \sum_{0}^{d(a)-1-\lambda} \omega^{\lambda+1+x} \cdot C_{\lambda+1+x}(a) + \sum_{0}^{\lambda} \omega^{x} \cdot C_{x}(a) + \omega^{\lambda} \text{ by (13)}$$
$$= S_{\lambda} f(a)$$

The theorem follows by $\mathbf{U}_1$.

The only successor function in terms of which the component functions are defined is S_0. This fact together with $C_{\mu}(0) = 0$ shows that the component functions only take values among the natural numbers. The degree function also only has values among the natural numbers. Hence the above theorem shows that every ordinal α less than ω^{ω} can be uniquely expressed in the form

$$\alpha = \omega^{\alpha_1} \cdot a_1 + \omega^{\alpha_2} \cdot a_2 + \dots + \omega^{\alpha_k} \cdot a_k$$

where $a_1, a_2, \dots, a_k$ are natural numbers and $\alpha_1, \alpha_2, \dots, \alpha_k$ is a decreasing sequence of ordinal numbers.

The ordinal α given above can be written as

$$S_{\alpha_k}^{a_k} S_{\alpha_{k-1}}^{a_{k-1}} \dots S_{\alpha_1}^{a_1} 0$$

where $S_{\alpha_i}^{a_i}$ is an abbreviation for $\underbrace{S_{\alpha_i} S_{\alpha_i} \dots S_{\alpha_i}}_{a_i}$

When this is done computation with ordinals written in normal form can be performed by successive applications of the rules involving successor and predecessor functions and other arithmetical functions, e.g.

$$\begin{aligned}
(\omega^3 + \omega^2 \cdot 2 + \omega + 3) + (\omega^2 + 1) &= S_0S_0S_0S_1S_2S_2S_30 + S_0S_20 \\
&= S_0(S_0S_0S_0S_1S_2S_2S_30 + S_20) \\
&= S_0[S_2(S_0S_0S_0S_1S_2S_2S_30 + 0)] \\
&= S_0S_2S_0S_0S_0S_1S_2S_2S_30 \\
&= S_0S_2S_2S_2S_30 \text{ by application of axiom } \mathbf{A} \\
&= \omega^3 + \omega^2 \cdot 3 + 1. \\
(\omega^2 + \omega \cdot 3) \cdot (\omega^3 + 1) &= S_1S_1S_1S_20 \cdot S_0S_30 \\
&= S_1S_1S_1S_20 \cdot 0 + S_1S_1S_1S_20 \cdot \omega^3 + S_1S_1S_1S_20 \\
&= \omega^{d(S_1S_1S_1S_20)+3} + S_1S_1S_1S_20 \\
&= \omega^5 + S_1S_1S_1S_20 \\
&= S_50 + S_1S_1S_1S_20 \\
&= S_1S_1S_1S_2S_50 \\
&= \omega^5 + \omega^2 + \omega \cdot 3
\end{aligned}$$

The addition defined above is not commutative. A new addition can therefore be defined by the following equation.

$$a \oplus b = b + a.$$

A countable number of functions T_{μ} are defined by the following equation.

$$T_{\mu} a = a \oplus \omega^{\mu}.$$

The following inference schema is proved.

$$\mathbf{U}_2 \qquad \frac{\begin{array}{r@{\;=\;}l} f(0) & g(0) \\ f(\mathsf{T}_\mu a) & \mathsf{H}_\mu(a, f(a)) \\ g(\mathsf{T}_\mu a) & \mathsf{H}_\mu(a, g(a)) \end{array}}{f(a) = g(a)}$$

In the following proof, functions are introduced some of whose arguments only take values among the natural numbers. The arithmetic of the natural numbers is used intuitively and proofs using transfinite induction on the natural numbers are permitted.

The function $\mathsf{G}^\mu_m(a,b)$ is defined by the following recursion.

$$\begin{aligned} \mathsf{G}^\mu_0(a,b) &= b \\ \mathsf{G}^\mu_{\mathsf{S}_0 m}(a,b) &= \mathsf{H}_\mu(\omega^\mu \cdot m + a, \mathsf{G}^\mu_m(a,b)) \end{aligned}$$

μ and m are restricted to the natural numbers.

$$\mathsf{G}^\mu_m(a, f(a)) = f(\omega^\mu \cdot m + a)$$

This is now proved.

$$\begin{aligned} \mathsf{G}^\mu_0(a, f(a)) &= f(a) \\ f(\omega^\mu \cdot 0 + a) &= f(a) \\ \mathsf{G}^\mu_{\mathsf{S}_0 m}(a, f(a)) &= \mathsf{H}_\mu(\omega^\mu \cdot m + a, \mathsf{G}^\mu_m(a, f(a))) \\ f(\omega^\mu \cdot \mathsf{S}_0 m + a) &= f((\omega^\mu + \omega^\mu \cdot m) + a) \\ &= f(\omega^\mu + (\omega^\mu \cdot m + a)) \\ &= f(\mathsf{T}_\mu(\omega^\mu \cdot m + a)) \\ &= \mathsf{H}_\mu(\omega^\mu \cdot m + a, f(\omega^\mu \cdot m + a)) \end{aligned}$$

The result follows by $\mathbf{U}_1$.

The inference schema is proved by induction on the degree of a.

For finite a $\mathsf{T}_0 a = \mathsf{S}_0 a$. Also $f(0) = g(0)$. The result is therefore true for $\mathsf{d}(a) = 0$. Suppose $f(a) = g(a)$ when $\mathsf{d}(a) < n$. Choose b so that

$$\begin{aligned} \mathsf{d}(b) &= \mathsf{d}(a) + 1 \\ b &= \sum_0^{\mathsf{d}(a)+1} \omega^x \cdot \mathsf{C}_x(b) \\ &= \omega^{\mathsf{d}(a)+1} \cdot \mathsf{C}_{\mathsf{d}(a)+1}(b) + \sum_0^{\mathsf{d}(a)} \omega^x \cdot \mathsf{C}_x(b) \end{aligned}$$

$$f(b) = \mathsf{G}^{\mathsf{d}(a)+1}_{\mathsf{C}_{\mathsf{d}(a)+1}(b)}\left(\sum_0^{\mathsf{d}(a)} \omega^x \cdot \mathsf{C}_x(b), f\left(\sum_0^{\mathsf{d}(a)} \omega^x \cdot \mathsf{C}_x(b)\right)\right)$$

by the result just proved

$$f\left(\sum_0^{\mathsf{d}(a)} \omega^x \cdot \mathsf{C}_x(b)\right) = g\left(\sum_0^{\mathsf{d}(a)} \omega^x \cdot \mathsf{C}_x(b)\right)$$

by the inductive assumption. Hence $f(b) = g(b)$ and the schema is proved.

A number of results involving subtraction are now proved.

(16)
$$\begin{aligned}
\mathsf{P}_\mu a \dot- b &= \mathsf{P}_\mu(a \dot- b)\\
\mathsf{P}_\mu a \dot- 0 &= \mathsf{P}_\mu a\\
\mathsf{P}_\mu(a \dot- 0) &= \mathsf{P}_\mu a\\
\mathsf{P}_\mu a \dot- \mathsf{S}_\nu b &= \mathsf{P}_\nu(\mathsf{P}_\mu a \dot- b)\\
\mathsf{P}_\mu(a \dot- \mathsf{S}_\nu b) &= \mathsf{P}_\mu \mathsf{P}_\nu(a \dot- b)\\
&= \mathsf{P}_\nu \mathsf{P}_\mu(a \dot- b) \text{ by (1)}
\end{aligned}$$

(17)
$$\begin{aligned}
(a \dot- b) \dot- c &= (a \dot- c) \dot- b\\
(a \dot- b) \dot- 0 &= a \dot- b\\
(a \dot- 0) \dot- b &= a \dot- b\\
(a \dot- b) \dot- \mathsf{S}_\mu c &= \mathsf{P}_\mu[(a \dot- b) \dot- c]\\
(a \dot- \mathsf{S}_\mu c) \dot- b &= \mathsf{P}_\mu(a \dot- c) \dot- b\\
&= \mathsf{P}_\mu[(a \dot- c) \dot- b] \text{ by (17)}
\end{aligned}$$

(18)
$$\begin{aligned}
a \dot- a &= 0\\
0 \dot- 0 &= 0\\
\mathsf{S}_\mu a \dot- \mathsf{S}_\mu a &= \mathsf{P}_\mu(\mathsf{S}_\mu a \dot- a)\\
&= \mathsf{P}_\mu \mathsf{S}_\mu a \dot- a \text{ by (17)}\\
\text{Let } f(a) &= \mathsf{P}_\mu \mathsf{S}_\mu a \dot- a\\
f(0) &= \mathsf{P}_\mu \mathsf{S}_\mu 0 \dot- 0 = 0 \dot- 0 = 0\\
\text{If } \nu < \mu,\ f(\mathsf{S}_\nu a) &= \mathsf{P}_\mu \mathsf{S}_\mu \mathsf{S}_\nu a \dot- \mathsf{S}_\nu a\\
&= \mathsf{P}_\nu(\mathsf{P}_\mu \mathsf{S}_\mu a \dot- a)\\
&= \mathsf{P}_\nu \mathsf{P}_\mu \mathsf{S}_\mu a \dot- a \text{ by (17)}\\
&= \mathsf{P}_\mu \mathsf{S}_\mu a \dot- a\\
&= f(a)\\
\text{If } \nu = \mu,\ f(\mathsf{S}_\nu a) &= \mathsf{P}_\mu \mathsf{S}_\mu \mathsf{S}_\mu a \dot- \mathsf{S}_\mu a\\
&= \mathsf{P}_\mu \mathsf{P}_\mu \mathsf{S}_\mu \mathsf{S}_\mu a \dot- a\\
&= \mathsf{P}_\mu \mathsf{S}_\mu a \dot- a \text{ by (vi)}\\
&= f(a)\\
\text{If } \nu > \mu,\ f(\mathsf{S}_\nu a) &= \mathsf{P}_\mu \mathsf{S}_\mu \mathsf{S}_\nu a \dot- \mathsf{S}_\nu a\\
&= \mathsf{S}_\nu a \dot- \mathsf{S}_\nu a \text{ by (vi).}
\end{aligned}$$

Hence we can prove $\mathsf{S}_\mu a \dot- \mathsf{S}_\mu a = 0$ if we can prove $\mathsf{S}_\nu a \dot- \mathsf{S}_\nu a = 0$ for all sufficiently large ν. Choose $\nu > \mathsf{d}(a)$.

$$\begin{aligned}
\text{Then } \mathsf{S}_\nu a &= \omega^\nu\\
\omega^\nu \dot- \omega^\nu &= \mathsf{P}_\nu \mathsf{S}_\nu 0 = 0.
\end{aligned}$$

The following result is sometimes useful.

(19)
$$\begin{aligned}
\mathsf{T}_\mu \mathsf{S}_\nu a &= \mathsf{S}_\nu \mathsf{T}_\mu a\\
\mathsf{T}_\mu \mathsf{S}_\nu a &= \omega^\mu + (a + \omega^\nu)\\
&= (\omega^\mu + a) + \omega^\nu\\
&= \mathsf{S}_\nu \mathsf{T}_\mu a.
\end{aligned}$$

The Difference Function

$$|a,b| = (a \dot- b) + (b \dot- a).$$

The following schema holds.

$$\frac{|a,b| = 0}{a = b}$$

Before proving this scheme the following result is proved.

(20) $$a \dot{-} b = 0 \text{ if } \mathsf{d}(b) \geq \mathsf{d}(a)$$

If $b = 0$ $\mathsf{d}(b) = 0$ and the result holds vacuously

$$a \dot{-} \mathsf{S}_\mu b = \mathsf{P}_\mu(a \dot{-} b)$$
$$\mathsf{P}_\mu 0 = 0.$$

The schema is now proved.

If $(a \dot{-} b) + (b \dot{-} a) = 0$
Suppose $\mathsf{d}(a) \geq \mathsf{d}(b)$
Then $(a \dot{-} b) + (b \dot{-} a) = a \dot{-} b$
Suppose $\mathsf{d}(b) \geq \mathsf{d}(a)$
Then $(a \dot{-} b) + (b \dot{-} a) = b \dot{-} a$.

We may therefore suppose in general

$$a \dot{-} b = 0 \text{ and } \mathsf{d}(a) > \mathsf{d}(b)$$

By Cantor's Normal Form theorem

$$a = \mathsf{S}_0^{n_0}\mathsf{S}_1^{n_1} \ldots \mathsf{S}_{\mathsf{d}(a)}^{n_{\mathsf{d}(a)}} 0$$
$$b = \mathsf{S}_0^{m_0}\mathsf{S}_1^{m_1} \ldots \mathsf{S}_{\mathsf{d}(b)}^{n_{\mathsf{d}(b)}} 0$$

where $n_{\mathsf{d}(a)} > 0$, $n_{\mathsf{d}(b)} > 0$ and $n_i \geq 0$ for $i < \mathsf{d}(a)$
and $m_i \geq 0$ for $i < \mathsf{d}(b)$

$$\begin{aligned}\text{Hence } a \dot{-} b &= \mathsf{P}_0^{m_0}\mathsf{P}_1^{m_1} \ldots \mathsf{P}_{\mathsf{d}(b)}^{m_{\mathsf{d}(b)}} \mathsf{S}_0^{n_0} \ldots \mathsf{S}_{\mathsf{d}(b)}^{n_{\mathsf{d}(a)}} 0 \\ &= \mathsf{P}_{\mathsf{d}(b)}^{m_{\mathsf{d}(b)}} \mathsf{S}_0^{n_0} \ldots \mathsf{S}_{\mathsf{d}(a)}^{n_{\mathsf{d}(a)}} 0 \text{ by (16)} \\ &= \mathsf{P}_{\mathsf{d}(b)}^{m_{\mathsf{d}(b)}} \mathsf{S}_{\mathsf{d}(b)+1}^{m_{\mathsf{d}(b)+1}} \ldots \mathsf{S}_{\mathsf{d}(a)}^{n_{\mathsf{d}(a)}} 0 \\ &\neq 0 \;\; \mathsf{d}(a) > \mathsf{d}(b).\end{aligned}$$

This is a contradiction. We may, therefore, suppose

$$\mathsf{d}(a) = \mathsf{d}(b).$$

Suppose $\mathsf{C}_{\mathsf{d}(a)}(a) \neq \mathsf{C}_{\mathsf{d}(a)}(b)$.
We may suppose $\mathsf{C}_{\mathsf{d}(a)}(b) < \mathsf{C}_{\mathsf{d}(a)}(a)$

$$\begin{aligned}\text{Then } a \dot{-} b &= \mathsf{P}_{\mathsf{d}(a)}^{\mathsf{C}_{\mathsf{d}(a)}(b)} \mathsf{S}_{\mathsf{d}(a)}^{\mathsf{C}_{\mathsf{d}(a)}(a)} 0 \\ &= \mathsf{S}_{\mathsf{d}(a)}^{\mathsf{C}_{\mathsf{d}(a)}(a) - \mathsf{C}_{\mathsf{d}(a)}(b)} 0 \\ &\neq 0\end{aligned}$$

Hence $C_{d(a)}(a) = C_{d(a)}(b)$
We may next prove $C_{d(a)-1}(a) = C_{d(a)-1}(b)$
and in general $C_i(a) = C_i(b)$ $i \leq d(a)$
Hence $a = b$.

An extension of the formalisation to ordinals greater than ω^ω

The ordinals less than ω^ω can be represented using successor functions indexed by the natural numbers. In the development of the arithmetic it is necessary to use some of the arithmetic of the natural numbers used in the indexing. By taking more successor functions and using indices extending into infinite ordinals it is possible to extend this formalisation to ordinals greater than ω^ω. It is necessary, however, to use some of the arithmetic of the indexing infinite ordinals. If the preceding formalisation of ordinals less than ω^ω is accepted it is then possible to consider successor functions indexed by such ordinals and to formalise ordinal arithmetic for ordinals less than ω^{ω^ω}. This procedure can, of course, be repeated and even greater ordinals considered.

REFERENCES

[1] Goodstein, R. L., "Logic-free formalisations of recursive arithmetic," *Mathematica Scandinavia,* v. 2 (1954), pp. 247-261.

[2] Vučković, V., "Partially ordered recursive arithmetics," *Mathematica Scandinavia,* v. 7 (1959), pp. 305-320.

University of Leicester
Leicester, England