Prime Factorisation and the Ordinals Below ε_0

The Schütte-Matula correspondence, with a short proof and references

Orientation

The ordinal ε_0 is not the largest ordinal. It is the *first epsilon number*, the least fixed point of ordinal exponentiation with base ω. It is countable, and every ordinal strictly below it has a finite hereditary Cantor normal form.

$$\varepsilon_0 = \min\left\{\xi > 0 : \omega^{\xi} = \xi\right\} = \sup_{k<\omega} \omega_k, \omega_0 = 1, \omega_{k+1} = \omega^{\omega_k}$$

The coding

Let p_m be the m-th prime, with $p_1 = 2$, $p_2 = 3$, and so on. Write each non-zero $\alpha < \varepsilon_0$ in Cantor normal form, allowing repeated terms:

$$\alpha = \omega^{\alpha_1} + \cdots + \omega^{\alpha_r}, \quad \alpha_1 \geq \cdots \geq \alpha_r$$

Define its positive-integer code recursively by

$$C(0) = 1, \qquad C(\alpha) = \prod_{j=1}^{r} p_{C(\alpha_j)}$$

Equivalently, if $\alpha = \omega^{\beta} + \gamma$ with $\gamma < \omega^{\beta}$, then

$$C(\omega^{\beta} + \gamma) = p_{C(\beta)} C(\gamma)$$

For a correspondence with $\mathbb{N}_0 = \{0,1,2,\ldots\}$ rather than the positive integers, simply map α to C(α) - 1.

The inverse map

For a positive integer n, list its prime factors with multiplicity as $n = p_{m_1} \cdots p_{m_r}$. Define recursively

$$O(1) = 0, \qquad O(n) = \omega^{O(m_1)} \# \cdots \# \omega^{O(m_r)}$$

Here # is the natural (Hessenberg) sum. For these monomials it simply collects equal powers of ω. Equivalently, sort the exponents $O(m_j)$ in decreasing order and use ordinary ordinal addition. In prime-power notation:

$$n = \prod_{m \geq 1} p_m^{a_m}, \qquad O(n) = \mathop{\#}_{m : a_m > 0} \omega^{O(m)} a_m$$

Examples

Integer	Factorisation	Ordinal
1	empty product	0
2	p_1	1
3	p_2	ω
4	$p_1{}^2$	2
5	p_3	ω^ω
6	$p_1 p_2$	ω + 1
7	p_4	ω^2
9	$p_2{}^2$	ω · 2

Thus 5 first becomes the prime index 3, and then 3 is decoded recursively: $O(5) = \omega^{O(3)} = \omega^\omega$. The exponents in the ordinary prime factorisation become finite Cantor coefficients; the genuinely recursive part is the decoding of the prime indices.

Why it is a bijection

The recursion terminates. In the ordinal direction, every exponent α_j is a proper subterm of the hereditary Cantor normal form. In the integer direction, if p_m divides $n > 1$, then $m < n$. Unique prime factorisation recovers exactly the multiset of encoded exponents, while uniqueness of Cantor normal form recovers the ordinal. Structural induction therefore gives

$$O(C(\alpha)) = \alpha \ \ (\alpha < \varepsilon_0), \qquad C(O(n)) = n \ \ (n \geq 1)$$

Where the construction appears

Matula's 1968 note gives the prime-index recursion as a bijection between positive integers and finite rooted trees [1]. Göbel gave a later proof and enumeration [2]. Cantor normal forms identify ordinals below ε_0 with finite rooted non-plane trees: the immediate subtrees of the root encode the exponents $\alpha_1,\ldots,\alpha_r$. Combining these two identifications yields precisely the correspondence above.

The ordinal formulation is explicit in Weiermann [4, pp. 193-194], who calls it a Matula-number or Schütte-style coding. Weiermann and Woods give the recurrence displayed above and say that it occurs in Schütte's proof-theory book [3] and is also known as Matula-coding [5, Remark 1].

References

[1] D. W. Matula, “A natural rooted tree enumeration by prime factorization,” *SIAM Review* 10 (1968), 273. Author's abstract.

[2] F. Göbel, “On a 1-1-correspondence between rooted trees and natural numbers,” *Journal of Combinatorial Theory, Series B* 29(1) (1980), 141-143. doi:10.1016/0095-8956(80)90049-0.

[3] K. Schütte, *Proof Theory*, Grundlehren der mathematischen Wissenschaften 225, Springer, 1977. doi:10.1007/978-3-642-66473-1.

[4] A. Weiermann, “Analytic combinatorics, proof-theoretic ordinals, and phase transitions for independence results,” *Annals of Pure and Applied Logic* 136(1-2) (2005), 189-218. doi:10.1016/j.apal.2005.05.012. See pp. 193-194 for the explicit isomorphism.

[5] A. Weiermann and A. R. Woods, “Some natural zero one laws for ordinals below ε_0,” in *How the World Computes*, LNCS 7318, Springer (2012), 723-732. doi:10.1007/978-3-642-30870-3_73; open-access record. Remark 1 gives the recurrence and the name “Matula-coding.”

Terminology varies between *Matula numbers*, *Matula coding*, and *Schütte-style coding*. The first name foregrounds the integer-rooted-tree bijection; the latter two foreground its use as an ordinal notation.