Formalising Erdős and Larson: Ordinal Partition Theory

Lawrence C Paulson FRS

Computer Laboratory

UNIVERSITY OF CAMBRIDGE

Leeds Logic Seminar, 14/03/2023

Supported by the ERC Advanced Grant ALEXANDRIA (742178).

Steps towards the formalisation of mathematics

- *Euclid*: unifying Greek geometry under an axiomatic system
- *Cauchy, Weierstrass*: removing infinitesimals from analysis (and more)
- *Dedekind, Cantor, Frege, Zermelo*: set theory and the axiom of choice
- *Whitehead, Russell, Bourbaki*: formal (or super-rigorous) mathematics
- *de Bruijn*: AUTOMATH, a type theory for mathematics

Now it's widely accepted that essentially
<u>all</u> mathematics is formalisable

The development of simple type theory

- Whitehead and Russell's ramified types for *Principia Mathematica*
- Simplified by Ramsey; formalised by Church
- First implemented on a computer by Michael JC Gordon
- Sophisticated implementations today include HOL Light and Isabelle/HOL

But is formalised maths possible?

Whitehead and Russell needed 362 pages to prove 1+1=2!

We have better formal systems than theirs.

Gödel proved that all reasonable formal systems must be incomplete!

But mathematicians also work from axioms!

Church proved that first-order logic is undecidable!

We want to **assist** people, not to **replace** them.

De Bruijn foresaw difficulties back in 1968!

As to the question what part of mathematics can be written in AUTOMATH, it should first be remarked that we do not possess a workable definition of the word "mathematics".

Quite often a mathematician jumps from his mathematical language into a kind of metalanguage, obtains results there, and uses these results in his original context. It seems to be very hard to create a single language in which such things can be done without any restriction.

And yet, in practice things seem to go okay.

Mathematics in Isabelle/HOL

- Jordan curve theorem
- Central limit theorem
- Residue theorem
- Prime number theorem
- Gödel's incompleteness theorems
- Algebraic closure of a field
- Verification of the Kepler conjecture*

- Matrix theory, e.g. Perron–Frobenius
- Analytic number theory, eg Hermite–Lindemann
- Nonstandard analysis
- Homology theory
- Topology
- Complex roots via Sturm sequences
- Measure, integration and probability theory

Distinctive features of Isabelle/HOL

- Simple types with axiomatic type classes
- Powerful automation: proofs *and* counterexamples
- Structured proof language
- Interactive development environment (PIDE)
- User-definable mathematical notation
- "Literate" proof documents can be generated in LaTeX
- An archive of over ***600 proof developments; 385 authors*** and nearly ***3 million lines of code***

Can we do *set theory* in higher-order logic?

- HOL is actually weaker than **Zermelo** set theory
- … but we can simply add a type of ZF sets with the usual axioms.
- [our framework presupposes the *axiom of choice*]
- … and develop cardinals, ordinal arithmetic, order types and the rest.

Partition notation: $\alpha \longrightarrow (\beta, \gamma)^n$

$[A]^n$ denotes the set of unordered n-element sets of elements of A

if $[\alpha]^n$ is partitioned ("coloured") into two parts (0, 1) then there's either

- a subset $B \subseteq \alpha$ of order type β whose n-sets are all coloured by 0
- a subset $C \subseteq \alpha$ of order type γ whose n-sets are all coloured by 1

Infinite Ramsey theorem: $\omega \longrightarrow (\omega, \omega)^n$

Erdős' problem (for 2-element sets)

$\alpha \longrightarrow (\alpha, 2)$ is trivial

$\alpha \longrightarrow (|\alpha| + 1, \omega)$ **fails** for $\alpha > \omega$

So which countable ordinals α satisfy $\alpha \longrightarrow (\alpha, 3)$?

It turns out that α must be a power of ω

In 1987, Erdős offered a $1000 prize for a full solution

Material formalised for this project

$\omega^2 \longrightarrow (\omega^2, m)$ (Specker)

$\omega^{1+\alpha n} \longrightarrow (\omega^{1+\alpha}, 2^n)$ (Erdős and Milner)

$\omega^\omega \longrightarrow (\omega^\omega, m)$ (Milner, Larson)

Plus background theories: Cantor normal form for ordinals; facts about order types; the Nash-Williams partition theorem

Project done with Mirna Džamonja and Angeliki Koutsoukou-Argyraki

Paul Erdős and E. C. Milner, 1972

$\omega^{1+\alpha n} \longrightarrow (\omega^{1+\alpha}, 2^n)$ for α an ordinal and n a natural number

"We have known this result since 1959"
(it's in Milner's 1962 PhD thesis)

It's a five-page paper that needed a **full-page correction** in 1974.

We now conclude the proof of the theorem.

Since β is denumerable and nonzero, there is a sequence $(\gamma_n : n < \omega)$ which repeats each element of B infinitely often, i.e. such that

$$(11) \qquad |\{n : \gamma_n = \nu\}| = \aleph_0 \qquad (\nu \in B).$$

Since tp $S = \alpha\beta$, we may write $S = S^{(0)} = \bigcup (\nu \in B) A_\nu^{(0)}(<)$.

Let $n < \omega$ and suppose we have already chosen elements $x_i \in S (i < n)$ and a subset

$$(12) \qquad S^{(n)} = \bigcup (\nu \in B) A_\nu^{(n)}(<)$$

of S of order type $\alpha\beta$. Since α is right-SI, $A_{\gamma_n}^{(n)}$ contains a final section A' such that $A_{\gamma_n}^{(n)} \cap \{x_0, \ldots, x_{n-1}\} < A'$. By (10), there are $x_n \in A'$, a strictly increasing map $g_n : B \to B$ and sets $A_\nu^{(n+1)} (\nu \in B)$ such that

$$(13) \qquad g_n(\gamma_i) = \gamma_i \qquad (i \leq n),$$

$$(14) \qquad A_\nu^{(n+1)} \subset K_1(x_n) \cap A_{g_n(\nu)}^{(n)} \qquad (\nu \in B).$$

From the definition of A', it follows that

$$(15) \qquad x_n \in A_{\gamma_n}^{(n)} \subset S^{(n)}$$

and

$$(16) \qquad x_i < x_n \quad \text{if} \quad i < n \quad \text{and} \quad x_i \in A_{\gamma_n}^{(n)}.$$

$S^{(n+1)}$ is defined by equation (12) with n replaced by $n+1$. It follows by induction that there are x_n, $A_\nu^{(n)} (\nu \in B)$, $S^{(n)}$ and g_n such that (12)–(16) hold for $n < \omega$.

Let $Z = \{x_n : n < \omega\}$. If $m < n < \omega$, then by (15), (14), and (12) we have that

Key steps of Erdős and Milner's proof

- Every ordinal is a "strong type" *(about 200 lines of machine proof)*
- A "remark" about indecomposable ordinals *(72 lines)*
- A key lemma: $\alpha\beta \longrightarrow (\min(\gamma, \omega\beta), 2k)$ if $\alpha \longrightarrow (\gamma, k)$ for $k \geq 2$ *(about 960 lines)*
- The main theorem $\omega^{1+\alpha n} \longrightarrow (\omega^{1+\alpha}, 2^n)$ by induction on n *(about 30 lines)*
- Larson's corollary: $\omega^{nk} \longrightarrow (\omega^n, k)$ *(about 35 lines)*

Every ordinal is a “strong type”

We will say that β is a strong type if, whenever tp $B{=}\beta$ and $D \subset B$, then there are $n{<}\omega$ and sets $D_1, \ldots, D_n \subset D$ such that

(5) tp D_i *is*

(6) *if* $M \subset$

```
proposition strong_ordertype_eq:
  assumes D: "D ⊆ elts β" and "Ord β"
  obtains L where "⋃(List.set L) = D" "⋀X. X ∈ List.set L ⟹ indecomposable (tp X)"
    and "⋀M. ⟦M ⊆ D; ⋀X. X ∈ List.set L ⟹ tp (M ∩ X) ≥ tp X⟧ ⟹ tp M = tp D"
proof -
  define φ where "φ ≡ inv_into D (ordermap D VWF)"
  then have bij_φ: "bij_betw φ (elts (tp D)) D"
    using D bij_betw_inv_into down ordermap_bij by blast
  have φ_cancel_left: "⋀d. d ∈ D ⟹ φ (ordermap D VWF d) = d"
    by (metis D φ_def bij_betw_inv_into_left down_raw ordermap_bij small_iff_range total_on_
  have φ_cancel_right: "⋀γ. γ ∈ elts (tp D) ⟹ ordermap D VWF (φ γ) = γ"
    by (metis φ_def f_inv_into_f ordermap_surj subsetD)
  have "small D" "D ⊆ ON"
    using assms down elts_subset_ON [of β] by auto
  then have φ_less_iff: "⋀γ δ. ⟦γ ∈ elts (tp D); δ ∈ elts (tp D)⟧ ⟹ φ γ < φ δ ⟷ γ < δ"
    using ordermap_mono_less [of _ _ VWF D] bij_betw_apply [OF bij_φ] VWF_iff_Ord_less φ_can
    by (metis ON_imp_Ord Ord_linear2 less_V_def order.asym)
```

A remark about indecomposable ordinals

```
proposition indecomposable_imp_Ex_less_sets:
  assumes indec: "indecomposable α" and "α > 1" and A: "tp A = α" "small A" "A ⊆ ON"
    and "x ∈ A" and A1: "tp A1 = α" "A1 ⊆ A"
  obtains A2 where "tp A2 = α" "A2 ⊆ A1" "{x} ≪ A2"
proof -
  have "Ord α"
    using indec indecomposable_imp_Ord by blast
  have "Limit α"
    by (simp add: assms indecomposable_imp_Limit)
  define φ where "φ ≡ inv_into A (ordermap A VWF)"
  then have bij_φ: "bij_betw φ (elts α) A"
    using A bij_betw_inv_into down ordermap_bij by blast
  have bij_om: "bij_betw (ordermap A VWF) A (elts α)"
    using A down ordermap_bij by blast
  define γ where "γ ≡ ordermap A VWF x"
  have γ: "γ ∈ elts α"
    unfolding γ_def using A ‹x ∈ A› down by auto
  then have "Ord γ"
    using Ord_in_Ord ‹Ord α› by blast
  define B where "B ≡ φ ` (elts (succ γ))"
  show thesis
  proof
    have "small A1"
      by (meson ‹small A› A1 smaller_than_small)
    then have "tp (A1 - B) ≤ tp A1"
```

If $x \in A$ and $A_1 \subseteq A$, with type $A, A_1 = \alpha$, then there is $A_2 \subseteq A_1$ such that $\{x\} < A_2$.

$\alpha\beta \longrightarrow (\min(\gamma, \omega\beta), 2k)$ if $\alpha \longrightarrow (\gamma, k)$

- Assume there is no $X \in [\alpha\beta]^{2k}$ such that $[X]^2$ is 1-coloured
- Assume there is no $C \subseteq \alpha\beta$ of order type γ such that $[C]^2$ is 0-coloured
- Then show there is a $Z \subseteq \alpha\beta$ of order type $\omega\beta$ such that $[Z]^2$ is 0-coloured

this will require generating an ω-chain of sets of type β

```
theorem Erdos_Milner_aux:
  assumes part: "partn_lst_VWF α [ord_of_nat k, γ] 2"
    and indec: "indecomposable α" and "k > 1" "Ord γ" and β: "β ∈ elts ω1"
  shows "partn_lst_VWF (α*β) [ord_of_nat (2*k), min γ (ω*β)] 2"
proof (cases "α=1 ∨ β=0")
  case True
  show ?thesis
  proof (cases "β=0")
    case True
    moreover have "min γ 0 = 0"
      by (simp add: min_def)
    ultimately show ?thesis
      by (simp add: partn_lst_triv0 [where i=1])
  next
    case False
    then obtain "α=1" "Ord β"
      by (meson ON_imp_Ord Ord_ω1 True β elts_subset_ON)
    then obtain i where "i < Suc (Suc 0)" "[ord_of_nat k, γ] ! i ≤ α"
      using partn_lst_VWF_nontriv [OF part] by auto
    then have "γ ≤ 1"
      using ‹α=1› ‹k > 1› by (fastforce simp: less_Suc_eq)
    then have "min γ (ω*β) ≤ 1"
      by (metis Ord_1 Ord_ω Ord_linear_le Ord_mult ‹Ord β› min_def order_trans)
    moreover have "elts β ≠ {}"
      using False by auto
    ultimately show ?thesis
      by (auto simp: True ‹Ord β› ‹β≠0› ‹α=1› intro!: partn_lst_triv1 [where i=1])
  qed
next
  case False
  then have "α ≠ 1" "β ≠ 0"
    by auto
```

$$\alpha\beta \longrightarrow (\min(\gamma, \omega\beta), 2k)$$

$$\text{if } \alpha \longrightarrow (\gamma, k)$$

Equation (8) with its one-line proof

(8) *If A ⊂ S, then there is X ∈ [A*

This follows from the hypothesi

```
have Ak0: "∃X ∈ [A]⇗k⇖. f ` [X]⇗2⇖ ⊆ {0}" —‹remark (8) about @{term"A ⊆ S"}›
  if A_αβ: "A ⊆ elts (α*β)" and ot: "tp A ≥ α" for A
proof -
  let ?g = "inv_into A (ordermap A VWF)"
  have "small A"
    using down that by auto
  then have inj_g: "inj_on ?g (elts α)"
    by (meson inj_on_inv_into less_eq_V_def ordermap_surj ot subset_trans)
  have Aless: "⋀x y. ⟦x ∈ A; y ∈ A; x < y⟧ ⟹ (x,y) ∈ VWF"
    by (meson Ord_in_Ord VWF_iff_Ord_less ‹Ord(α*β)› subsetD that(1))
  then have om_A_less: "⋀x y. ⟦x ∈ A; y ∈ A; x < y⟧ ⟹ ordermap A VWF x < ordermap A VWF y
    by (auto simp: ‹small A› ordermap_mono_less)
  have α_sub: "elts α ⊆ ordermap A VWF ` A"
    by (metis ‹small A› elts_of_set less_eq_V_def ordertype_def ot replacement)
  have g: "?g ∈ elts α → elts (α * β)"
    by (meson A_αβ Pi_I' α_sub inv_into_into subset_eq)
  then have fg: "f ∘ (λX. ?g ` X) ∈ [elts α]⇗2⇖ → {..<2}"
    by (rule nsets_compose_image_funcset [OF f _ inj_g])
  have g_less: "?g x < ?g y" if "x < y" "x ∈ elts α" "y ∈ elts α" for x y
    using Pi_mem [OF g]
    by (meson A_αβ Ord_in_Ord Ord_not_le ord ‹small A› dual_order.trans elts_subset_ON inv_
  obtain i H where "i < 2" "H ⊆ elts α"
    and ot_eq: "tp H = [k,γ]!i" "(f ∘ (λX. ?g ` X)) ` (nsets H 2) ⊆ {i}"
    using ii partn_lst_E [OF part fg] by (auto simp: eval_nat_numeral)
  then consider (0) "i=0" | (1) "i=1"
    by linarith
  then show ?thesis
  proof cases
    case 0
    then have "f ` [inv_into A (ordermap A VWF) ` H]⇗2⇖ ⊆ {0}"
      using ot_eq ‹H ⊆ elts α› α_sub by (auto simp: nsets_def [of _ k] inj_on_inv_into elim
    moreover have "finite H ∧ card H = k"
```

```
theorem Erdos_Milner:
  assumes ν: "ν ∈ elts ω1"
  shows "partn_lst_VWF (ω↑(1 + ν * ord_of_nat n)) [ord_of_nat (2^n), ω↑(1+ν)] 2"
proof (induction n)
  case 0
  then show ?case
    using partn_lst_VWF_degenerate [of 1 2] by simp
next
  case (Suc n)
  have "Ord ν"
    using Ord_ω1 Ord_in_Ord assms by blast
  have "1+ν ≤ ν+1"
    by (simp add: ‹Ord ν› one_V_def plus_Ord_le)
  then have [simp]: "min (ω ↑ (1 + ν)) (ω * ω ↑ ν) = ω ↑ (1+ν)"
    by (simp add: ‹Ord ν› oexp_add min_def)
  have ind: "indecomposable (ω ↑ (1 + ν * ord_of_nat n))"
    by (simp add: ‹Ord ν› indecomposable_ω_power)
  show ?case
  proof (cases "n = 0")
    case True
    then show ?thesis
      using partn_lst_VWF_ω_2
  next
    case False
    then have "Suc 0 < 2 ^ n"
      using less_2_cases not_less_eq by fastforce
    then have "partn_lst_VWF (ω ↑ (1 + ν * n) * ω ↑ ν) [ord_of_nat (2 * 2 ^ n), ω ↑ (1 + ν)] 2"
      using Erdos_Milner_aux [OF Suc ind, where β = "ω↑ν"] ‹Ord ν› ν
      by (auto simp: countable_oexp)
    then show ?thesis
      using ‹Ord ν› by (simp add: mult_succ mult.assoc oexp_add)
  qed
qed
```

Main Theorem

Suppose (2) holds for some integer $h \geq 1$. Applying the above theorem with $k=2^h$, $\alpha=\omega^{1+\nu h}$, $\beta=\omega^\nu$, $\gamma=\omega^{1+\nu}$, we see that (2) also holds with h replaced by $h+1$. Since (2) holds trivially for $h=1$, it follows that (2) holds for all $h<\omega$.

Jean Larson, 1973

$\omega^\omega \longrightarrow (\omega^\omega, m)$ for m a natural number

Proved by CC Chang in a 56-page paper (*J. Combinatorial Theory* A) and generalised by EC Milner

Simplified by Larson to 17 pages, including a new proof of $\omega^2 \longrightarrow (\omega^2, m)$

A few key definitions

Work with *finite increasing sequences*

- $W(n) = \{(a_0, a_1, \ldots, a_{n-1}) : a_0 < a_1 < \cdots < a_{n-1} < \omega\}$ has order type ω^n
- $W = W(0) \cup W(1) \cup W(2) \cup \cdots$ has order type ω^ω

Given $f : [W]^2 \to \{0,1\}$ such that there is no $M \in [W]^m$ s.t. $[M]^2$ is 1-coloured

Show there is a $X \subseteq W$ of order type ω^ω such that $[X]^2$ is 0-coloured

Interaction schemes

For $x, y \in W$, write $x = a_1 * a_2 * \cdots * a_k\,(*a_{k+1})$ and $y = b_1 * b_2 * \cdots * b_k$

put $c = (|a_1|, |a_1| + |a_2|, \ldots, |a_1| + |a_2| + \cdots + |a_k|\,(+ |a_{k+1}|))$

define $i(\{x, y\}) = c * a_1 * d * b_1 * a_2 * b_2 * \cdots * a_k * b_k\,(*a_{k+1})$

(this classifies how consecutive segments in x, y interact)

By Erdős–Milner we can assume $|x| < |y|$

The Nash-Williams partition theorem

A set $A \subseteq W$ is *thin* if for all $s, t \in A$, the sequence s is not an initial segment of t.

Given an infinite set $M \subseteq \omega$, a thin set A, a function $h : \{s \in A : s \subseteq M\} \to \{0,1\}$.

Then there exists an $i \in \{0,1\}$ and an infinite set $N \subseteq M$ so that $h(\{s \in A : s \subseteq N\}) \subseteq \{i\}$.

The three main lemmas

Lemma 3.6. *For every function* $g : [W]^2 \to \{0, 1\}$, *there exists an infinite set* $N \subseteq \omega$ *and a sequence* $\{j_k : k < \omega\}$, *so that for any* $k < \omega$ *with* $k > 0$, *and any pair* $\{x, y\}$ *of form* k *with* $(n_k) < i(\{x, y\}) \subseteq N$, $g(\{x, y\}) = j_k$.

150 lines, using Nash-Williams

Lemma 3.7. *For every infinite set* N *and every* $m, l < \omega$ *with* $l > 0$, *there is an* m *element set* M, *so that for every* $\{x, y\} \subseteq M$, $\{x, y\}$ *has form* l *and* $i(\{x, y\}) \subseteq N$.

900 lines, including inductive definitions of sequences

Lemma 3.8. *For any infinite set* $N \subseteq \omega$ *there is a set* $X \subseteq W$ *of type* ω^ω *so that for any pair* $\{x, y\} \subseteq X$, *there is an* $l < \omega$, *so that* $\{x, y\}$ *is of form* l *and if* $l > 0$, *then* $(n_l) < i(\{x, y\}) \subseteq N$.

1700 lines: more sequences and an order type calculation

… and the main theorem

Now we finish the proof of Theorem 3.1 using these three lemmas. First we apply Lemma 3.6 to f and obtain an infinite set N and a sequence $\{j_k : k < \omega\}$. Then for each $k < \omega$ with $k > 0$, we apply Lemma 3.7 to k, m and $\{n_l : k < l < \omega\}$ and obtain an m element set M_k, so that for any $\{x, y\} \subset M_k$, $f(\{x, y\}) = j_k$. Thus we may conclude that for any $k < \omega$ with $k > 0$, $j_k = 0$. Next we apply Lemma 3.8 to N and obtain a set $X \subseteq W$ of type ω^ω, so that for any $\{x, y\} \subseteq X$, there is an $l < \omega$ for which $\{x, y\}$ has form l and if $l > 0$, then $(n_l) < i(\{x, y\}) \subseteq N$. Thus on pairs $\{x, y\} \subseteq X$ which are not of form 0, $f(\{x, y\}) = j_l = 0$ for some l. By assumption, for any pair $\{x, y\}$ of form 0, $f(\{x, y\}) = 0$, so $f([X]^2) = \{0\}$, and the theorem follows.

150 lines

Why are these machine proofs so long?

- *The level of detail* in published proofs varies **immensely**
- … plus my lack of expertise in the area
- “Obvious” claims—about order types, cardinality, combinatorial intuitions— don’t have obvious proofs
- And some of the constructions are gruesome

This sort of inductive definition is tricky!

Let $d^1 = (n_1, n_2, \dots, n_{k+1}) = (d_1^1, d_2^1, \dots, d_{k+1}^1)$ and let a_1^1 be the sequence of the first d_1^1 elements of N greater than d_{k+1}^1. Now suppose we have constructed $d^1, a_1^1, \dots, d^i, a_1^i$. Let $d^{i+1} = (d_1^{i+1}, \dots, d_{k+1}^{i+1})$ be the first $k + 1$ elements of N greater than the last element of a_1^i, and let a_1^{i+1} be the first d_1^{i+1} elements of N greater than d_{k+1}^{i+1}. This defines $d^1, d^2, \dots, d^m, a_1^1, a_1^2, \dots, a_1^m$. Let the rest of the sequences be defined in the order that follows, so that for any i and j, a_j^i is the sequence of the least $(d_j^i - d_{j-1}^i)$ elements of N all of which are larger than the largest element of the sequence previously defined:

$$(a_1^m)a_2^1, a_2^2, a_2^3, \dots, a_2^m, a_3^1, \dots, a_3^m, \dots, a_k^1, \dots, a_k^m, a_{k+1}^m, a_{k+1}^{m-1}, \dots, a_{k+1}^1.$$

Other formalisations within ALEXANDRIA

- Transcendence of Certain Infinite Series (criteria by Hančl and Rucki)
- Irrationality Criteria for Series by Erdős and Straus
- Irrational Rapidly Convergent Series (a theorem by J. Hančl)
- Counting Complex Roots
- Budan–Fourier Theorem and Counting Real Roots
- Localization of a Commutative Ring
- Projective Geometry
- Quantum Computation and Information
- **Grothendieck Schemes**

Brief remarks on Grothendieck Schemes

- Build-up of mainstream structures in algebraic geometry: presheaves and sheaves of rings, locally ringed spaces, affine schemes
- the *spectrum of a ring* is a locally ringed space, hence an affine scheme
- any *affine scheme* is a scheme
- *They said it couldn't be done* in simple type theory.
- But we did it faster and with less manpower than the Lean guys.
- One key technique: a structuring mechanism known as *locales*.*
- *led by Anthony Bordg*

What can mathematicians expect from proof technology in the future?

- Ever-growing libraries of definitions and theorems
- … with advanced search
- Verification of dull but necessary facts
- … and exhibiting counterexamples
- Detection of analogous developments, with hints for proof steps
- Warnings of simple omissions, e.g. "doesn't S need to be compact?"
- *A careful and increasingly intelligent assistant*