Towards Solid Abelian Groups: A Formal Proof of Nöbeling's Theorem

Dagur Asgeirsson ☑ �� ⑩ University of Copenhagen, Denmark

- Abstract

Condensed mathematics, developed by Clausen and Scholze over the last few years, is a new way of studying the interplay between algebra and geometry. It replaces the concept of a topological space by a more sophisticated but better-behaved idea, namely that of a condensed set. Central to the theory are solid abelian groups and liquid vector spaces, analogues of complete topological groups.

Nöbeling's theorem, a surprising result from the 1960s about the structure of the abelian group of continuous maps from a profinite space to the integers, is a crucial ingredient in the theory of solid abelian groups; without it one cannot give any nonzero examples of solid abelian groups. We discuss a recently completed formalisation of this result in the Lean theorem prover, and give a more detailed proof than those previously available in the literature. The proof is somewhat unusual in that it requires induction over ordinals – a technique which has not previously been used to a great extent in formalised mathematics.

2012 ACM Subject Classification General and reference → Verification; Computing methodologies \rightarrow Representation of mathematical objects; Mathematics of computing \rightarrow Mathematical software

Keywords and phrases Condensed mathematics, Nöbeling's theorem, Lean, Mathlib, Interactive theorem proving

Digital Object Identifier 10.4230/LIPIcs.ITP.2024.6

Supplementary Material *Software*: https://github.com/leanprover-community/mathlib4/ blob/ba9f2e5baab51310883778e1ea3b48772581521c/Mathlib/Topology/Category/Profinite/ Nobeling.lean archived at swh:1:cnt:2fb2985994d43409a52761d0e853d37deeabdc74

Funding Dagur Asgeirsson: The author was supported by the Danish National Research Foundation (DNRF) through the "Copenhagen Center for Geometry and Topology" under grant no. DNRF151.

Acknowledgements First and foremost, I would like to thank Johan Commelin for encouraging me to start seriously working on this project when we were both in Banff attending the workshop on formalisation of cohomology theories last year. I had useful discussions related to this work with Johan, Kevin Buzzard, Adam Topaz, and Dustin Clausen. I am indebted to all four of them for providing helpful feedback on earlier drafts of this paper. Any project formalising serious mathematics in Lean depends on Mathlib, and this one is no exception. I am grateful to the Lean community as a whole for building and maintaining such a useful mathematical library, as well as providing an excellent forum for Lean-related discussions through the Zulip chat. Finally, I would like to thank the anonymous referees for helpful feedback.

1 Introduction

Nöbeling's theorem says that the abelian group $C(S,\mathbb{Z})$ of continuous maps from a profinite space S to the integers, is a free abelian group. In fact, the original statement [14, Satz 1] is the more general result that bounded maps from any set to the integers form a free abelian group, but this special case has recently been applied [16, Theorem 5.4] in the new field of condensed mathematics (see also [15, 5, 3]).

We report on a recently completed formalisation of this theorem using the Lean 4 theorem prover [12], building on its Mathlib library of formalised mathematics (which was recently ported from Lean 3, see [11, 10]). The proof uses the well-ordering principle and a tricky induction over ordinals. This is the first use of the induction principle for ordinals in Mathlib outside the directory containing the theory of ordinals. Often, one can replace such transfinite constructions by appeals to Zorn's lemma. The author is not aware of any proof of Nöbeling's theorem that does this, or otherwise avoids induction over ordinals¹.

When formalising a nontrivial proof, one inevitably makes an effort to organise the argument carefully. One purpose of this paper is to give a well-organised and detailed proof of Nöbeling's theorem, written in conventional mathematical language, which is essentially a by-product of the formalisation effort. This will hopefully be a more accessible proof than those that already exist in the literature; the one in [14] is in German, while the proofs of the result in [7, 16] are the same argument as the one presented here, but in significantly less detail. This is the content of section 4; some mathematical prerequisites are found in section 3.

In section 2 we give more details about the connection to condensed mathematics and in sections 5 and 6 we discuss the formalisation process and the integration into Mathlib.

Throughout the text, we use the symbol "Z" for external links, usually directly to the source code for the corresponding theorems and definitions in Mathlib. In order for the links to stay usable, they are all to a fixed commit to the master branch (the most recent one at the time of writing).

Mathlib is a growing library of mathematics formalised in Lean. All material is maintained continuously by a team of experts. There is a big emphasis on unity, meaning that there is one official definition of every concept, and it is the job of contributors to provide proofs that alternative definitions are equivalent. All the code in this project has been integrated into Mathlib; a process that took quite some time, as high standards are demanded of code that enters the library. However, it is an important part of formalisation to get the code into Mathlib, because doing so means that it stays usable to others in the future.

2 Motivation

Condensed mathematics [16, 15, 5] is a new theory developed by Clausen and Scholze (and independently by Barwick and Haine, who called the theory $pyknotic\ sets$ [3]). It has the purpose of generalising topology in a way that gives better categorical properties, which is desirable e.g. when the objects have both a topological and an algebraic structure. Condensed objects² can be described as sheaves on a certain site of profinite spaces. A topological abelian group A can be regarded as a condensed abelian group with S-valued points C(S,A) for profinite spaces S. Discrete abelian groups such as $\mathbb Z$ are important examples of topological abelian groups. There is a useful characterisation of discrete condensed sets (which leads to the same characterisation for more general condensed objects such as condensed abelian groups), which has been formalised in Lean 3 by the author in [1].

The discreteness characterisation can be stated somewhat informally as follows: A condensed set X is discrete if and only if for every profinite space $S = \varprojlim_i S_i$ (written as a cofiltered limit of finite discrete spaces), the natural map

$$\varinjlim_{i} X(S_i) \to X(S)$$

is an isomorphism.

¹ The proof does not use any ordinal arithmetic. However, it crucially uses the principle of induction over ordinals with a case split between limit ordinals and successor ordinals.

² This notion was first formalised in the *Liquid Tensor Experiment* [6, 17], see section 6 for a more detailed discussion.

There is a notion of completeness of condensed abelian groups, called being *solid* [16, Definition 5.1]. For the convenience of the reader, we give the informal definition here in Definition 1. First, we need to recall two facts about condensed abelian groups:

- The category of condensed abelian groups has all limits.
- The forgetful functor from condensed abelian groups to condensed sets has a left adjoint, denoted by $\mathbb{Z}[-]$ (adopted from the analogous relationship between the category of sets and the category of abelian groups).

▶ **Definition 1.** Let $S = \varprojlim_i S_i$ be a profinite space and define a condensed abelian group as follows:

$$\mathbb{Z}[S]^{\blacksquare} := \varprojlim_{i} \mathbb{Z}[S_{i}]$$

There is a natural map $\mathbb{Z}[S] \to \mathbb{Z}[S]^{\blacksquare}$, and we say that a condensed abelian group A is solid if for every profinite space S and every morphism $f: \mathbb{Z}[S] \to A$ of condensed abelian groups, there is a unique morphism $g: \mathbb{Z}[S]^{\blacksquare} \to A$ making the obvious triangle commute³.

Using the discreteness characterisation and Nöbeling's theorem, one can prove that for every profinite space S, there is a set I and an isomorphism of condensed abelian groups

$$\mathbb{Z}[S]^{\blacksquare} \cong \prod_{i \in I} \mathbb{Z}.$$

This structural result is essential to developing the theory of solid abelian groups. Without it one cannot even prove the existence of a nontrivial solid abelian group.

Since the proof of Nöbeling's theorem has nothing to do with condensed mathematics, people studying the theory might be tempted to skip the proof and use Nöbeling's theorem as a black box. Now that it has been formalised, they can do this with a better conscience. On the other hand, people interested in understanding the proof might want to turn to sections 3 and 4 of this paper for a more detailed account.

3 Preliminaries

For ease of reference, we collect in this section some prerequisites for the proof of Nöbeling's theorem. Most of them were already in Mathlib.

3.1 Order theory

- ▶ Remark 4. Taking the union of two sets is an example of a join operation.
- ▶ **Definition 5.** 🗹 A category C is filtered if it satisfies the following three conditions
 (i) C is nonempty.

 $^{^3\,}$ This definition has also been formalised by the author in Lean 3 in [2]

- (ii) For all objects X, Y, there exists an object Z and morphisms $f: X \to Z$ and $g: Y \to Z$.
- (iii) For all objects X, Y and all morphisms $f, g: X \to Y$, there exists an object Z and a morphism $h: Y \to Z$ such that $h \circ f = h \circ g$.

A category is cofiltered if the opposite category is filtered.

- ▶ Remark 6. A poset is filtered if and only if it is nonempty and directed.
- ▶ Remark 7. The poset of finite subsets of a given set is filtered.

3.2 Linear Independence

- ▶ Lemma 9. 🗹 Suppose we have a commutative diagram

where N, M, P are modules over a ring R, the top row is exact, and the bottom maps are the inclusion maps. If v and w are linearly independent, then u is linearly independent.

3.3 Cantor's intersection theorem

- ▶ Remark 11. Cantor's intersection theorem is often stated only for the special case of decreasing nested sequences of nonempty compact, closed subsets. The generalisation above can be proved by slightly modifying the standard proof of that special case.

3.4 Cofiltered limits of profinite spaces

- ▶ **Definition 12.** ∠ A profinite space is a totally disconnected compact Hausdorff space.
- ▶ Lemma 13. 🗹 Every profinite space has a basis of clopen subsets.
- ▶ Lemma 14. ☐ Every profinite space is totally separated, i.e. any two distinct points can be separated by clopen neighbourhoods.
- ▶ Remark 15. A topological space is profinite if and only if it can be written as a cofiltered limit of finite discrete spaces. See section 6 for a further discussion.
- ▶ Lemma 16. Any continuous map from a cofiltered limit of profinite spaces to a discrete space factors through one of the components.
- ▶ Remark 17. In particular, a continuous map from a profinite space

$$S = \varprojlim_i S_i$$

to a discrete space factors through one of the finite quotients S_i .

4 The theorem

This section is devoted to proving

We can immediately reduce this to proving Lemma 19 below as follows: Let I denote the set of clopen subsets of S. Then the map

$$S \to \prod_{i \in I} \{0, 1\}$$

whose i-th projection is given by the indicator function of the clopen subset i is a closed embedding.

▶ Lemma 19. 🗹 Let I be a set and let S be a closed subset of $\prod_{i \in I} \{0,1\}$. Then $C(S,\mathbb{Z})$ is a free abelian group.

To prove Lemma 19, we need to construct a basis of $C(S, \mathbb{Z})$. Our proposed basis is defined as follows:

- \blacksquare Choose a well-ordering on I.
- Let $e_{S,i} \in C(S,\mathbb{Z})$ denote the composition

$$S \longrightarrow \prod_{i \in I} \{0,1\} \xrightarrow{p_i} \{0,1\} \longrightarrow \mathbb{Z}$$

where p_i denotes the *i*-th projection map, and the other two maps are the obvious inclusions.

- Let P denote the set of finite, strictly decreasing sequences in I. Order these lexicographically.
- Let $ev_S: P \to C(S, \mathbb{Z})$ denote the map

$$(i_1, \cdots, i_r) \mapsto e_{S,i_1} \cdots e_{S,i_r}.$$

For $p \in P$, let $\Sigma_S(p)$ denote the span in $C(S,\mathbb{Z})$ of the set

$$\operatorname{ev}_S \left(\{ q \in P \mid q$$

Let E(S) denote the subset of P consisting of those elements whose evaluation cannot be written as a linear combination of evaluations of smaller elements of P, i.e.

$$E(S) := \{ p \in P \mid \operatorname{ev}_S(p) \notin \Sigma_S(p) \}.$$

In Subsection 4.2 we prove that the set $\operatorname{ev}_S(E(S))$ spans $C(S,\mathbb{Z})$, and in Subsection 4.3 we prove that the family

$$\operatorname{ev}_S: E(S) \to C(S, \mathbb{Z})$$

is linearly independent, concluding the proof of Nöbeling's theorem. Subsection 4.1 defines some notation which will be convenient for bookkeeping in the subsequent proof.

4.1 Notation and generalities

For a subset J of I we denote by

$$\pi_J: \prod_{i \in I} \{0,1\} \to \prod_{i \in I} \{0,1\}$$

the map whose *i*-th projection is p_i if $i \in J$, and 0 otherwise. These maps are continuous, and since source and target are compact Hausdorff spaces, they are also closed. Given a subset $S \subseteq \prod_{i \in I} \{0,1\}$, we let

$$S_J := \pi_J(S).$$

We can regard I with its well-ordering as an ordinal. Then I is the set of all strictly smaller ordinals. Given an ordinal μ , we let

$$\pi_{\mu} := \pi_{\{i \in I \mid i < \mu\}}$$

and

$$S_{\mu} := S_{\{i \in I \mid i < \mu\}}.$$

These maps induce injective \mathbb{Z} -linear maps

$$\pi_J^*: C(S_J, \mathbb{Z}) \to C(S, \mathbb{Z})$$

by precomposition.

Recall that we have defined P as the set of finite, strictly decreasing sequences in I, ordered lexicographically. We will use this notation throughout the proof of Nöbeling's theorem.

▶ Lemma 20. \square For $p \in P$ and $x \in S$, we have

$$\operatorname{ev}_S(p)(x) = \begin{cases} 1 & \text{if } \forall i \in p, \ x_i = 1\\ 0 & \text{otherwise.} \end{cases}$$

Proof. Obvious.

▶ Lemma 21. 🗹 Let J be a subset of I and let $p \in P$ be such that $i \in p$ implies $i \in J$. Then $\pi_J^*(ev_{S_J}(p)) = ev_S(p)$.

Proof. Since $i \in p$ implies $i \in J$, we have

$$x_i = \pi_I^*(x)_i$$

for all $x \in S$ and $i \in p$. The result now follows from Lemma 20.

- ▶ Remark 22. 🗹 The hypothesis in Lemma 21 holds in particular if $p \in E(S_J)$. Indeed, suppose $i \in p$, then if $i \notin J$, we have $\text{ev}_{S_J}(p) = 0$.
- ▶ **Lemma 23.** \square If μ' , μ are ordinals satisfying $\mu' < \mu$, then $E(S_{\mu'}) \subseteq E(S_{\mu})$.

Proof. Let $p \in E(S_{\mu'})$. Then every entry of p is $< \mu'$, and it suffices to show that if

$$ev_{S_{n}}(p) = \pi_{n'}^*(ev_{S_{n'}}(p))$$

is in the span of

$$\operatorname{ev}_{S_{\mu}} (\{q \in P \mid q < p\}) = \pi_{\mu'}^* \left(\operatorname{ev}_{S_{\mu'}} (\{q \in P \mid q < p\}) \right)$$

then $\operatorname{ev}_{S_{\mu'}}(p)$ is in the span of $\operatorname{ev}_{S_{\mu'}}(\{q \in P \mid q < p\})$. This follows by injectivity of $\pi_{\mu'}^*$.

4.2 Span

The following series of lemmas proves that $\operatorname{ev}_S(E(S))$ spans $C(S, \mathbb{Z})$.

ightharpoonup Lemma 24. The set P is well-ordered.

Proof sketch. Suppose not. Take a strictly decreasing sequence (p_n) in P. Let a_n denote first term of p_n . Then (a_n) is a decreasing sequence in I and hence eventually constant. Denote its limit by a. Let $q_n = p_n \setminus a_n$. Then there exists an N such that $(q_n)_{n \geq N}$ is a strictly decreasing sequence in P and we can repeat the process of taking the indices of the first factors, get a decreasing sequence in I whose limit is strictly smaller than a. Continuing this way, we get a strictly decreasing sequence in I, a contradiction.

- ▶ Remark 25. The proof sketch of Lemma 24 above is ill-suited for formalisation. Kim Morrison gave a formalised proof ∠, following similar ideas to those above, which used close to 300 lines of code. A few days later, Junyan Xu found a proof ∠ that was ten times shorter, directly using the inductive datatype WellFounded. This is the only result whose proof indicated in this paper differs significantly from the one used in the formalisation.
- ▶ Lemma 26. \square If $ev_S(P)$ spans $C(S,\mathbb{Z})$, then $ev_S(E(S))$ spans $C(S,\mathbb{Z})$.
- **Proof.** It suffices to show that $\operatorname{ev}_S(P)$ is contained in the span of $\operatorname{ev}_S(E(S))$. Suppose it is not, and let p be the smallest element of P whose evaluation is not in the span of $\operatorname{ev}_S(E(S))$ (this p exists by Lemma 24). Write $\operatorname{ev}_S(p)$ as a linear combination of evaluations of strictly smaller elements of P. By minimality of p, each term of the linear combination is in the span of $\operatorname{ev}_S(E(S))$, implying that p is as well, a contradiction.

Proof sketch. Since S is compact and the limit is Hausdorff, it suffices to show that the natural map from S to the limit of F induced by the projection maps $\pi_J: S \to S_J$ is bijective.

For injectivity, let $a, b \in S$ such that $\pi_J(a) = \pi_J(b)$ for all finite subsets J of I. For all $i \in I$ we have $a_i = \pi_{\{i\}}(a) = \pi_{\{i\}}(b) = b_i$, hence a = b.

For surjectivity, let $b \in \lim F$. Denote by

$$f_J: \lim F \to S_J$$

the projection maps. We need to construct an element a of C such that $\pi_J(a) = f_J(b)$ for all J. In other words, we need to show that the intersection

$$\bigcap_{J} \pi_{J}^{-1} \{ f_{J}(b) \},$$

where J runs over all finite subsets of I, is nonempty. By Cantor's intersection theorem 10, it suffices to show that this family is directed (all the fibres are closed by continuity of the π_J , and closed subsets of a compact Hausdorff space are compact). To show that it is directed, it suffices to show that for $J \subseteq K$, we have

$$\pi_K^{-1}\{f_K(b)\} \subseteq \pi_J^{-1}\{f_J(b)\}$$

(by Lemma 3). This follows easily because the transition maps in the limits are just restrictions of the π_J .

▶ **Lemma 28.** $\ \ \, \square$ Let J be a finite subset of I. Then $\operatorname{ev}_{S_J}(E(S_J))$ spans $C(S_J, \mathbb{Z})$.

Proof. By lemma 26, it suffices to show that $\operatorname{ev}_{S_J}(P)$ spans. For $x \in S_J$, denote by f_x the map $S_J \to \mathbb{Z}$ given by the Kronecker delta $f_x(y) = \delta_{xy}$.

Since S_J is finite, the set of continuous maps is actually the set of all maps, and the maps f_x span $C(S_J, \mathbb{Z})$.

Now let $j_1 > \cdots > j_r$ be a decreasing enumeration of the elements of J. Let $x \in S_J$ and let e_i denote e_{S_J,j_i} if $x_{j_i} = 1$ and $(1 - e_{S_J,j_i})$ if $x_{j_i} = 0$. Then

$$f_{j_i} = \prod_{i=1}^r e_i$$

is in the span of $P(S_J)$, as desired.

▶ Lemma 29. \square P spans $C(S, \mathbb{Z})$.

Proof. Let $f \in C(S, \mathbb{Z})$. Then by Lemmas 27 and 16, there is a $g \in C(S_J, \mathbb{Z})$ such that $f = \pi_J^*(g)$. Writing this g as a linear combination of elements of $E(S_J)$, by Lemma 21 we see that f is a linear combination of elements of P as desired.

4.3 Linear independence

▶ Notation 30. Regard I with its well-ordering as an ordinal. Let Q denote the following predicate on an ordinal $\mu \leq I$:

For all closed subsets S of $\prod_{i \in I} \{0, 1\}$, such that for all $x \in S$ and $i \in I$, $x_i = 1$ implies $i < \mu$, E(S) is linearly independent in $C(S, \mathbb{Z})$.

We want to prove the statement Q(I). We prove by induction on ordinals that $Q(\mu)$ holds for all ordinals $\mu \leq I$.

▶ **Lemma 31.** \square The base case of the induction, Q(0), holds.

Proof. In this case, S is empty or a singleton. If S is empty, the result is trivial. Suppose S is a singleton. We want to show that E(S) consists of only the empty list, which evaluates to 1 and is linearly independent in $C(S,\mathbb{Z}) \cong \mathbb{Z}$. Let $p \in P$ and suppose p is nonempty. Then it is strictly larger than the empty list. But the evaluation of the empty list is 1, which spans $C(S,\mathbb{Z}) \cong \mathbb{Z}$, and thus $\operatorname{ev}_S(p)$ is in the span of strictly smaller products, i.e. not in E(S).

4.3.1 Limit case

Let μ be a limit ordinal, S a closed subset such that for all $x \in S$ and $i \in I$, $x_i = 1$ implies $i < \mu$. In other words, $S = S_{\mu}$. Suppose $Q(\mu')$ holds for all $\mu' < \mu$. Then in particular $E(S_{\mu'})$ is linearly independent

$$\pi_{\mu'}^* \left(\operatorname{ev}_{S_{\mu'}} \left(\{ q \in P \mid q$$

Proof. If q < p, then every element of q is also $< \mu'$. Thus, by lemma 21,

$$\pi_{\mu'}^* \left(\operatorname{ev}_{S_{\mu'}}(q) \right) = \operatorname{ev}_{S_{\mu}}(q),$$

as desired.

▶ Lemma 33. 🗹

$$E(S_{\mu}) = \bigcup_{\mu' < \mu} E(S_{\mu'})$$

Proof. The inclusion from right to left follows from Lemma 23, so we only need to show that if $p \in E(S_{\mu})$ then there exists $\mu' < \mu$ such that $p \in E(S_{\mu'})$. Take μ' to be the supremum of the set $\{i+1 \mid i \in p\}$. Then $\mu' = 0$ if p is empty, and of the form i+1 for an ordinal $i < \mu$ if p is nonempty. In either case, $\mu' < \mu$.

Since every $i \in p$ satisfies $i < \mu' < \mu$, we have

$$\operatorname{ev}_{S_{\mu}}(p) = \pi_{\mu'}^*(\operatorname{ev}_{S_{\mu'}}(p))$$

and

$$\operatorname{ev}_{S_{\mu}} (\{q \in P \mid q < p\}) = \pi_{\mu'}^* \left(\operatorname{ev}_{S_{\mu'}} (\{q \in P \mid q < p\}) \right)$$

so if $ev_{S_{n'}}(p)$ is in the span of

$$\text{ev}_{S_{n'}} (\{ q \in P \mid q$$

then $ev_{S_n}(p)$ is in the span of

$$\operatorname{ev}_{S_n} (\{q \in P \mid q < p\}),$$

contradicting the fact that $p \in E(S_{\mu})$.

▶ Lemma 34. 🔼

$$\operatorname{ev}_{S_{\mu}}\left(E(S_{\mu})\right) = \bigcup_{\mu' < \mu} \pi_{\mu'}^{*}\left(\operatorname{ev}_{S_{\mu'}}\left(E(S_{\mu'})\right)\right)$$

Proof. This follows from a combination of Lemmas 21 and 33.

The family of subsets in the union in Lemma 34 is directed with respect to the subset relation (this follows from Lemmas 3, 23, and 21). The sets $\operatorname{ev}_{S_{\mu'}}(E(S_{\mu'}))$ are all linearly independent by the inductive hypothesis, and by injectivity of π^*_{μ} , their images under that map are as well. Thus, by Lemma 8, the union is linearly independent, and we are done.

4.3.2 Successor case

Let μ be an ordinal, S a closed subset such that for all $x \in S$ and $i \in I$, $x_i = 1$ implies $i < \mu + 1$. In other words, $S = S_{\mu+1}$. Suppose $Q(\mu)$ holds. Then in particular $\operatorname{ev}_{S_{\mu}} : E(S_{\mu}) \to C(S_{\mu}, \mathbb{Z})$ is linearly independent.

To prove the inductive step in the successor case, we construct a closed subset S' of $\prod_{i \in I} \{0,1\}$ such that for all $x \in S'$, $x_i = 1$ implies $i < \mu$, and a commutative diagram

$$0 \longrightarrow C(S_{\mu}, \mathbb{Z}) \xrightarrow{\pi_{\mu}^{*}} C(S, \mathbb{Z}) \xrightarrow{g} C(S', \mathbb{Z})$$

$$\stackrel{\operatorname{ev}_{S_{\mu}} \uparrow}{\longrightarrow} \stackrel{\operatorname{ev}_{S} \uparrow}{\longrightarrow} E(S) \longleftrightarrow E'(S)$$

$$(1)$$

where the top row is exact and E'(S) is the subset of E(S) consisting if those p with $\mu \in p$ (note that p necessarily starts with μ). For $p \in P$, we denote by $p^t \in P$ the sequence obtained by removing the first element of p (t stands for tail). The linear map p has the property that $g(\text{ev}_S(p)) = \text{ev}_{S'}(p^t)$ and $p^t \in E(S')$. Given such a construction, the successor step in the induction follows from lemma p.

► Construction 35. 🗹 Let

$$S_0 = \{ x \in S \mid x_\mu = 0 \},\$$

$$S_1 = \{ x \in S \mid x_\mu = 1 \},\$$

and

$$S' = S_0 \cap \pi_{\mu}(S_1).$$

Then S' satisfies the inductive hypothesis.

$$g_1^* - g_0^* : C(S, \mathbb{Z}) \to C(S', \mathbb{Z}),$$

which we denote by g.

▶ Lemma 37. 🗹 The top row in diagram (1) is exact.

Proof. We already know that π_{μ}^* is injective. Also, since $\pi_{\mu} \circ g_1 = \pi_{\mu} \circ g_0$, we have

$$g \circ \pi_{\mu}^* = 0.$$

Now suppose we have

$$f \in C(S, \mathbb{Z})$$
 with $g(f) = 0$.

We want to find an

$$f_{\mu} \in C(S_{\mu}, \mathbb{Z})$$
 with $f_{\mu} \circ \pi_{\mu} = f$.

Denote by

$$\pi'_{\mu}:\pi_{\mu}(S_1)\to S_1$$

the map that swaps the μ -th coordinate to 1. Since g(f) = 0, we have

$$f \circ g_1 = f \circ g_0$$

and hence the two continuous maps $f_{|S_0}$ and $f_{|S_1} \circ \pi'_{\mu}$ agree on the intersection

$$S' = S_0 \cap \pi_{\mu}(S_1)$$

Together, they define the desired continuous map f_{μ} on all of $S_0 \cup \pi_{\mu}(S_1) = S_{\mu}$.

▶ **Lemma 38.** \square If $p \in P$ starts with μ , then $g(ev_S(p)) = ev_{S'}(p^t)$.

Proof. This follows from considering all the cases given by Lemma 20. We omit the proof here and refer to the Lean proof linked above.

- ▶ Remark 39. If $p \in E(S)$ and $\mu \in p$, then p satisfies the hypotheses of Lemma 38.
- ▶ Lemma 40. \square If $p \in E(S)$ and $\mu \in p$, then $p^t \in E(S')$.

Proof. Contraposing the statement, it suffices to show that if

$$\operatorname{ev}_{S'}(p^t) \in \operatorname{Span}\left(\operatorname{ev}_{S'}\left(\left\{q \mid q < p^t\right\}\right)\right),$$

then

$$\operatorname{ev}_S(p) \in \operatorname{Span}\left(\operatorname{ev}_S\left(\{(q) \mid q < p\}\right)\right).$$

Given a $q \in P$ such that $i \in q$ implies $i < \mu$, we denote by $q^{\mu} \in P$ the sequence obtained by adding μ at the front. Write

$$g(\operatorname{ev}_S(p)) = \operatorname{ev}_{S'}(p^t) = \sum_{q < p^t} n_q \operatorname{ev}_{S'}(q) = \sum_{q < p^t} n_q g(\operatorname{ev}_S(q^\mu)).$$

Then by Lemma 37, there exists an $n \in C(S_{\mu}, \mathbb{Z})$ such that

$$\operatorname{ev}_{S}(p) = \pi_{\mu}^{*}(n) + \sum_{q < p^{t}} n_{q}(\operatorname{ev}_{S}(q^{\mu})).$$

Now it suffices to show that each of the two terms in the sum above is in the span of $\{\operatorname{ev}_S(q)\mid q< p\}$. The latter term is because $q< p^t$ implies $q^\mu< p$. The former term is because we can write n as a linear combination indexed by $E(S_\mu)$, and for $q\in E(S_\mu)$ we have $\pi_\mu^*\left(\operatorname{ev}_{S_\mu}(q)\right)=\operatorname{ev}_S(q)$ and $\mu\notin q$ so q< p.

▶ **Lemma 41.** \square The set E(S) is the disjoint union of $E(S_{\mu})$ and

$$E'(S) = \{ p \in E(S) \mid \mu \in p \}.$$

Proof. We already know by Lemma 23 that $E(S_{\mu}) \subseteq E(S)$. Also, as noted in Remark 22, if $p \in E(S_{\mu})$ then all elements of p are $< \mu$ and hence $p \notin E'(S)$. Now it suffices to show that if $p \in E(S) \setminus E'(S)$, then $p \in E(S_{\mu})$.

Since $p \in E(S) \setminus E'(S)$, every $i \in p$ satisfies $i < \mu$. We have

$$\operatorname{ev}_S(p) = \pi_u^*(\operatorname{ev}_{S_u}(p))$$

and

$$\operatorname{ev}_{S}(\{q \in P \mid q < p\}) = \pi_{\mu}^{*}(\operatorname{ev}_{S_{\mu}}(\{q \in P \mid q < p\}))$$

so if $ev_{S_u}(p)$ is in the span of

$$\text{ev}_{S_n} (\{ q \in P \mid q$$

then $ev_S(p)$ is in the span of

$$\operatorname{ev}_S (\{q \in P \mid q < p\}),$$

contradicting the fact that $p \in E(S)$.

The above lemmas prove all the claims made at the beginning of this section, concluding the inductive proof. \square

5 The formalisation

First a note on terminology: in the mathematical exposition of the proof in section 4, we have talked about continuous maps from S to \mathbb{Z} . Since \mathbb{Z} is discrete, these are the same as the locally constant maps. The statement we have formalised is Listing 1.

```
instance LocallyConstant.freeOfProfinite (S : Profinite.{u}) : Module.Free \mathbb Z (LocallyConstant S \mathbb Z)
```

Listing 1 Nöbeling's theorem

which says that the \mathbb{Z} -module of locally constant maps from S to \mathbb{Z} is free. When talking about locally constant maps, one does not have to specify a topology on the target, which is slightly more convenient when working in a proof assistant.

The actual proof is about closed subsets of the product $\prod_{i \in I} \{0, 1\}$, which is of course the same thing as the space of functions $I \to \{0, 1\}$. We implement it as the type I \to Bool, where Bool is the type with two elements called true and false. This is the canonical choice for a two-element discrete topological space in Mathlib.

5.1 The implementation of P and E(S)

We implemented the set P as the type Products I defined as

```
def Products (I : Type*) [LinearOrder I] := {1 : List I // 1.Chain' (.>.)}
```

The predicate 1.Chain' (\rightarrow) means that adjacent elements of the list 1 are related by ">". We define the evaluation ev_S of products as

```
def Products.eval (S : Set (I \rightarrow Bool)) (1 : Products I) : LocallyConstant S \mathbb{Z} := (1.val.map (e S)).prod
```

where 1.val.map (e S) is the list of $e_{S,i}$ for i in the list 1.val, and List.prod is the product of the elements of a list.

We define a predicate on Products

```
def Products.isGood (S : Set (I \rightarrow Bool)) (1 : Products I) : Prop := 1.eval S \notin Submodule.span \mathbb{Z} ((Products.eval S) '' {m | m < 1}) and then the set E(S) becomes def GoodProducts (S : Set (I \rightarrow Bool)) : Set (Products I) := {1 : Products I | 1.isGood S}
```

It is slightly painful to prove completely trivial lemmas like 20 and its corollary 21 in Lean. Indeed, these results are not mentioned in the proof of [16, Theorem 5.4]. Although trivial, they are used often in the proof of the theorem and hence very important to making the proof work. Reading an informal proof of this theorem, one might never realise that these trivialities are used. This is an example of a useful by-product of formalisation; more clarity of exposition.

5.2 Ordinal induction

When formalising an inductive proof of any kind, one has to be very precise about what statement one wants to prove by induction. This is almost never the case in traditional mathematics texts. For example, the proof of [16, Theorem 5.4] claims to be proving by

induction that E(S) is a basis of $C(S,\mathbb{Z})$, not just that it is linearly independent. Furthermore, the set I is not fixed throughout the inductive proof which makes it somewhat unclear what the inductive hypothesis actually says. Working inside the topological space $\prod_{i \in I} \{0,1\}$ for a fixed set I throughout the proof was convenient in the successor step. This avoided problems that are solved by abuse of notation in informal texts, such as regarding a set as the same thing as its image under a continuous embedding.

The statement of the induction principle for ordinals in Mathlib is the following⁴:

```
\begin{array}{lll} \textbf{def Ordinal.limitRecOn } \{ \texttt{Q} : \texttt{Ordinal} \rightarrow \texttt{Sort } \_ \} & (\texttt{o} : \texttt{Ordinal}) \\ & (\texttt{H}_1 : \texttt{Q} \ \texttt{O}) \\ & (\texttt{H}_2 : \forall \ \texttt{o}, \ \texttt{Q} \ \texttt{o} \rightarrow \ \texttt{Q} \ (\texttt{succ o})) \\ & (\texttt{H}_3 : \forall \ \texttt{o}, \ \texttt{IsLimit} \ \texttt{o} \rightarrow \ (\forall \ \texttt{o'} < \texttt{o}, \ \texttt{Q} \ \texttt{o'}) \rightarrow \ \texttt{Q} \ \texttt{o}) : \\ & \texttt{Q} \ \texttt{o} \end{array}
```

In our setting, given a map Q: Ordinal \rightarrow Prop (in other words, a *predicate on ordinals*)⁵, we can prove $Q(\mu)$ for any ordinal μ if three things hold:

- \blacksquare The zero case: Q(0) holds.
- The successor case: for all ordinals λ , $Q(\lambda)$ implies $Q(\lambda + 1)$.
- The limit case: for every limit ordinal λ , if $Q(\lambda')$ holds for every $\lambda' < \lambda$, then $Q(\lambda)$ holds. Finding the correct predicate Q on ordinals was essential to the success of this project:

The inequality

```
o \leq Ordinal.type (\cdot < \cdot : I 	o I 	o Prop)
```

means that $o \leq I$ when I is considered as an ordinal, and the proposition contained S o is defined as

```
def contained {I : Type*} [LinearOrder I] [IsWellOrder I (\cdot<·)] (S : Set (I \rightarrow Bool)) (o : Ordinal) : Prop := \forall f, f \in S \rightarrow \forall (i : I), f i = true \rightarrow ord I i < o
```

and ord I i is an abbreviation for

```
Ordinal.typein (\cdot \cdot \cdot : I \rightarrow I \rightarrow Prop) i
```

i.e. the element $i \in I$ considered as an ordinal. The conclusion

LinearIndependent \mathbb{Z} (GoodProducts.eval S)

means that the map $\operatorname{ev}_S: E(S) \to C(S, \mathbb{Z})$ is linearly independent.

As is often the case, this is quite an involved statement that we are proving by induction, and when writing informally, mathematicians wouldn't bother to specify the map Q: $Ordinal \rightarrow Prop$ explicitly.

⁴ We have altered the notation slightly to match the notation in this paper.

⁵ Prop is Sort 0

5.3 Piecewise defined locally constant maps

In the proof of Lemma 37, we defined a locally constant map $S_{\mu} \to \mathbb{Z}$ by giving locally constant maps from S_0 and $\pi_{\mu}(S_1)$ that agreed on the intersection, and noting that this gives a locally constant map from the union which is equal to S_{μ} . To do this in Lean, the following definition was added to Mathlib \mathbb{Z} :

```
def LocallyConstant.piecewise {X Z : Type*} [TopologicalSpace X] {C_1 C_2 : Set X} (h_1 : IsClosed C_1) (h_2 : IsClosed C_2) (h : C_1 \cup C_2 = Set.univ) (f : LocallyConstant C_1 Z) (g : LocallyConstant C_2 Z) (hfg : \forall (x : X) (hx : x \in C_1 \cap C_2), f (x, hx.1) = g (x, hx.2)) [\forall j, Decidable (j \in C_1)] : LocallyConstant X Z where toFun i := if hi : i \in C_1 then f (i, hi) else g (i, (compl_subset_iff_union.mpr h) hi) isLocallyConstant := omitted
```

It says that given locally constant maps f and g defined respectively on closed subsets C_1 and C_2 which together cover the space X, such that f and g agree on $C_1 \cap C_2$, we get a locally constant map defined on all of X. This seems like exactly what we need in the above-mentioned proof. However, there is a subtlety, in that because of how the rest of the inductive proof is structured, we want the sets S_0 , $\pi_{\mu}(S_1)$ and S_{μ} all to be considered as subsets of the underlying topological space $\prod_{i \in I} \{0,1\}$. To use LocallyConstant.piecewise, we would have to consider S_{μ} as the underlying topological space and S_0 and $\pi_{\mu}(S_1)$ as subsets of it. This is possible and is what was done initially, but a cleaner solution is to define a variant of LocallyConstant.piecewise:

```
def LocallyConstant.piecewise' {X Z : Type*} [TopologicalSpace X]
       \{C_0 \ C_1 \ C_2 : Set X\}
       (h_0 : C_0 \subseteq C_1 \cup C_2) (h_1 : IsClosed C_1) (h_2 : IsClosed C_2)
       (f_1 : LocallyConstant C_1 Z) (f_2 : LocallyConstant C_2 Z)
       [DecidablePred (\cdot \in C_1)]
       (hf : \forall x (hx : x \in C<sub>1</sub> \cap C<sub>2</sub>), f<sub>1</sub> \langlex, hx.1\rangle = f<sub>2</sub> \langlex, hx.2\rangle) :
       LocallyConstant C<sub>0</sub> Z
which satisfies the equations
lemma LocallyConstant.piecewise'_apply_left {X Z : Type*} [TopologicalSpace X]
        \{ \mathtt{C}_0 \ \mathtt{C}_1 \ \mathtt{C}_2 \ : \ \mathtt{Set} \ \mathtt{X} \} \ (\mathtt{h}_0 \ : \ \mathtt{C}_0 \ \subseteq \ \mathtt{C}_1 \ \cup \ \mathtt{C}_2 ) 
       (h_1 : IsClosed C_1) (h_2 : IsClosed C_2)
       (f_1: LocallyConstant C_1 Z) (f_2: LocallyConstant C_2 Z)
       [DecidablePred (\cdot \in C_1)]
       (hf : \forall x (hx : x \in C<sub>1</sub> \cap C<sub>2</sub>), f<sub>1</sub> \langlex, hx.1\rangle = f<sub>2</sub> \langlex, hx.2\rangle)
       (x : C_0) (hx : x.val \in C_1) :
       piecewise' h_0 h_1 h_2 f_1 f_2 hf x = f_1 \langle x.val, hx \rangle
and
lemma LocallyConstant.piecewise'_apply_right {X Z : Type*} [TopologicalSpace X]
       \{C_0 \ C_1 \ C_2 : Set \ X\} \ (h_0 : C_0 \subseteq C_1 \cup C_2)
       (h_1 : IsClosed C_1) (h_2 : IsClosed C_2)
       (f_1: LocallyConstant C_1 Z) (f_2: LocallyConstant C_2 Z)
       [DecidablePred (\cdot \in C_1)]
       (hf : \forall x (hx : x \in C<sub>1</sub> \cap C<sub>2</sub>), f<sub>1</sub> \langlex, hx.1\rangle = f<sub>2</sub> \langlex, hx.2\rangle)
       (x : C_0) (hx : x.val \in C_2) :
       piecewise' h_0 h_1 h_2 f_1 f_2 hf x = f_2 \langle x.val, hx \rangle
```

Here C_0, C_1 , and C_2 are subsets of the same underlying topological space X; C_1 and C_2 are closed sets covering C_0 , and f_1 and f_2 are locally constant maps defined on C_1 and C_2 respectively, such that f_1 and f_2 agree on the intersection. This fits the application perfectly and shortened the proof of Lemma 37 considerably. Subtleties like this come up frequently, and can stall the formalisation process, especially when formalising general topology. When formalising Gleason's theorem \Box (another result in general topology relevant to condensed mathematics, see [16, Definition 2.4] and [8]), similar subtleties arose about changing the "underlying topological space" to a subset of the previous underlying topological space.

The phenomenon that it is sometimes more convenient to formalise the definition of an object rather as a subobject of some bigger object is, of course, well known. It was noted in the context of group theory by Gonthier et al. during the formalisation of the odd order theorem, see [9, Section 3.3].

5.4 Reflections on the proof

The informal proof in [16] is about half a page; 21 lines of text. Depending on how one counts (i.e. what parts of the code count as part of the proof and not just prerequisites), the formalised proof is somewhere between 1500 and 3000 lines of Lean code. A big part of the difference is because of omissions in the proof in [16].

A more fair comparison would be with the entirety of section 4 in this paper, which is an account of all the mathematical contents of the formalised proof. Still, there is quite a big difference, which is mostly explained by the pedantry of proof assistants, as discussed in subsections 5.1, 5.2, and 5.3.

5.5 Mathlib integration

As discussed in recent papers by Nash [13] and Best et al. [4], when formalising mathematics in Lean, it is desirable to develop as much as possible directly against Mathlib. Otherwise, the code risks going stale and unusable, while if integrated into Mathlib it becomes part of a library that is continuously maintained.

The development of this project took place on a branch of Mathlib, all code being written in new files. This was a good workflow to get the formalisation done as quickly as possible, because if new code is put in the "correct places" immediately, one has to rebuild part of Mathlib to be able to use that code in other places, which can be a slow process if changes are made deep in the import hierarchy.

The proof of Nöbeling's theorem described in this paper has now been fully integrated into Mathlib. An unusually large portion of the code was of no independent interest, which resulted in a pull request adding one huge file, which Johan Commelin and Kevin Buzzard kindly reviewed in great detail, improving both the style and performance of the code.

6 Towards condensed mathematics in Mathlib

The history of condensed mathematics in Lean started with the *Liquid Tensor Experiment* (LTE) [6, 17]. This is an example of a formalisation project that was in some sense too big to be integrated into Mathlib. Nevertheless, it was a big success in that it demonstrated the capabilities of Lean and its community by fully formalising the complicated proof of a highly nontrivial theorem about so-called liquid modules. Moreover, it provided a setting in which to experiment with condensed mathematics and find the best way to do homological algebra in Lean. As mentioned above, the goal of LTE was to formalise one specialised theorem.

This is somewhat orthogonal to the goal of Mathlib which is to build a coherent, unified library of formalised mathematics. It is thus understandable that the contributors of LTE chose to focus on completing the task at hand instead of spending time on moving some parts of the code to Mathlib. Now that both LTE and the port of Mathlib to Lean 4 have been completed, we are seeing some important parts of LTE being integrated into Mathlib.

The definition **Z** of a condensed object was recently added to Mathlib. During a masterclass on formalisation of condensed mathematics organised in Copenhagen in June 2023, participants collaborated, under the guidance of Kevin Buzzard and Adam Topaz, on formalising as much condensed mathematics as possible in one week (all development took place in Lean 4 and the goal was to write material for Mathlib). The code can be found in the masterclass GitHub repository **Z** and much of it has already made it into Mathlib.

Profinite spaces form a rich category of topological spaces and there is more work other than Nöbeling's theorem to be done in Mathlib. Being the building blocks of condensed sets, it is important to develop a good API for profinite spaces in Mathlib. There, profinite spaces are defined as totally disconnected compact Hausdorff spaces. It is proved \checkmark that every profinite space can be expressed as a cofiltered limit (more precisely, over the poset of its discrete quotients). It is also proved \checkmark that the category of profinite spaces has all limits and that the forgetful functor to topological spaces preserves them \checkmark . From this we can extract the following useful theorem:

▶ **Theorem 42.** A topological space is profinite if and only if it can be written as a cofiltered limit of finite discrete spaces.

The story about profinite spaces as limits does not end there, though. Sometimes it is not enough to know just that some profinite space can be written as a limit, but rather that there is a specific limit formula for it. Lemma 27 gives one specific way of writing a compact subset of a product as a cofiltered limit, which can be useful. Another example can be extracted from [1]. This is the fact that the identity functor on the category of profinite spaces is right Kan extended from the inclusion functor from finite sets to profinite spaces along itself. This gives another limit formula for profinite spaces, coming from the limit formula for right Kan extensions, and is useful when formalising the definition of solid abelian groups [2].

It can also be useful to regard the category of profinite spaces as the pro-category of the category of finite sets. The definition of pro-categories and this equivalence of categories would make for a nice formalisation project and be a welcome contribution to Mathlib.

7 Conclusion and future work

By formalising Nöbeling's theorem, we have illustrated that the induction principle for ordinals in Mathlib can be used to prove nontrivial theorems outside the theory ordinals themselves. Another contribution is the detailed proof given in section 4, and of course as mentioned before, it is an important step for the formalisation of condensed mathematics to continue.

A natural next step in the formalisation of the theory of solid abelian groups is to port the code in [1, 2] to Lean 4 and get it into Mathlib. Then one can put together the discreteness characterisation and Nöbeling's theorem to prove the structural results about $\mathbb{Z}[S]^{\blacksquare}$, which would lead us one step closer to an example of a nontrivial solid abelian group in Mathlib.

More broadly, it is important to continue moving as much as possible of the existing Lean code about condensed mathematics (from the LTE and the Copenhagen masterclass) into Mathlib.

References

1 Dagur Asgeirsson. Formalising discrete condensed sets. https://github.com/dagurtomas/lean-solid/tree/discrete, 2023.

- 2 Dagur Asgeirsson. Formalising solid abelian groups. https://github.com/dagurtomas/lean-solid/, 2023.
- 3 Clark Barwick and Peter Haine. Pyknotic objects, i. basic notions, 2019. arXiv:1904.09966.
- 4 Alex J. Best, Christopher Birkbeck, Riccardo Brasca, and Eric Rodriguez Boidi. Fermat's Last Theorem for Regular Primes. In Adam Naumowicz and René Thiemann, editors, 14th International Conference on Interactive Theorem Proving (ITP 2023), volume 268 of Leibniz International Proceedings in Informatics (LIPIcs), pages 36:1–36:8, Dagstuhl, Germany, 2023. Schloss Dagstuhl Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ITP.2023.36.
- 5 Dustin Clausen and Peter Scholze. Condensed mathematics and complex geometry. https://people.mpim-bonn.mpg.de/scholze/Complex.pdf, 2022.
- 6 Johan Commelin, Adam Topaz et al. Liquid tensor experiment. https://github.com/leanprover-community/lean-liquid, 2022.
- 7 László Fuchs. Infinite Abelian Groups. ISSN. Elsevier Science, 1970.
- 8 Andrew M. Gleason. Projective topological spaces. *Illinois Journal of Mathematics*, 2(4A):482–489, 1958. doi:10.1215/ijm/1255454110.
- 9 Georges Gonthier, Andrea Asperti, Jeremy Avigad, Yves Bertot, Cyril Cohen, François Garillot, Stéphane Le Roux, Assia Mahboubi, Russell O'Connor, Sidi Ould Biha, Ioana Pasca, Laurence Rideau, Alexey Solovyev, Enrico Tassi, and Laurent Théry. A machine-checked proof of the odd order theorem. In Sandrine Blazy, Christine Paulin-Mohring, and David Pichardie, editors, Interactive Theorem Proving, pages 163–179, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
- 10 The mathlib Community. The lean mathematical library. In Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2020, pages 367–381, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3372885.3373824.
- 11 Leonardo de Moura, Soonho Kong, Jeremy Avigad, Floris van Doorn, and Jakob von Raumer. The lean theorem prover (system description). In Amy P. Felty and Aart Middeldorp, editors, *Automated Deduction CADE-25*, pages 378–388, Cham, 2015. Springer International Publishing.
- 12 Leonardo de Moura and Sebastian Ullrich. The lean 4 theorem prover and programming language. In André Platzer and Geoff Sutcliffe, editors, *Automated Deduction CADE 28*, pages 625–635, Cham, 2021. Springer International Publishing.
- Oliver Nash. A Formalisation of Gallagher's Ergodic Theorem. In Adam Naumowicz and René Thiemann, editors, 14th International Conference on Interactive Theorem Proving (ITP 2023), volume 268 of Leibniz International Proceedings in Informatics (LIPIcs), pages 23:1–23:16, Dagstuhl, Germany, 2023. Schloss Dagstuhl Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ITP.2023.23.
- 14 Georg Nöbeling. Verallgemeinerung eines Satzes von Herrn E. Specker. Invent. Math., 6:41–55, 1968. doi:10.1007/BF01389832.
- Peter Scholze. Lectures on analytic geometry. http://www.math.uni-bonn.de/people/scholze/Analytic.pdf, 2019.
- Peter Scholze. Lectures on condensed mathematics. https://www.math.uni-bonn.de/people/scholze/Condensed.pdf, 2019.
- 17 Peter Scholze. Liquid tensor experiment. Experimental Mathematics, 31(2):349–354, 2022. doi:10.1080/10586458.2021.1926016.