# A Formalisation of the Balog–Szemerédi–Gowers Theorem in Isabelle/HOL

Angeliki Koutsoukou-Argyraki
University of Cambridge
United Kingdom
ak2110@cam.ac.uk

Mantas Bakšys
University of Cambridge
United Kingdom
mb2412@cam.ac.uk

Chelsea Edmonds
University of Cambridge
United Kingdom
cle47@cl.cam.ac.uk

## Abstract

We describe our formalisation in the interactive theorem prover Isabelle/HOL of the Balog–Szemerédi–Gowers Theorem, a profound result in additive combinatorics which played a central role in Gowers's proof deriving the first effective bounds for Szemerédi's Theorem. The proof is of great mathematical interest given that it involves an interplay between different mathematical areas, namely applications of graph theory and probability theory to additive combinatorics involving algebraic objects. This interplay is what made the process of the formalisation, for which we had to develop formalisations of new background material in the aforementioned areas, more rich and technically challenging. We demonstrate how locales, Isabelle's module system, can be employed to handle such interplays in mathematical formalisations. To treat the graph-theoretic aspects of the proof, we make use of a new, more general undirected graph theory library developed by Edmonds, which is both flexible and extensible. In addition to the main theorem, which, following our source, is formulated for difference sets, we also give an alternative version for sumsets which required a formalisation of an auxiliary triangle inequality. We moreover formalise a few additional results in additive combinatorics that are not used in the proof of the main theorem. This is the first formalisation of the Balog–Szemerédi–Gowers Theorem in any proof assistant to our knowledge.

*CCS Concepts:* • **Mathematics of computing → Combinatorics**; *Graph theory*; • **Theory of computation → Logic and verification**; Automated reasoning.

***Keywords:*** interactive theorem proving, proof assistant, formalisation of mathematics, Isabelle/HOL, additive combinatorics, graph theory, probabilistic method.

*CPP '23, January 16–17, 2023, Boston, MA, USA*
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0026-2/23/01.
https://doi.org/10.1145/3573105.3575680

**ACM Reference Format:**
Angeliki Koutsoukou-Argyraki, Mantas Bakšys, and Chelsea Edmonds. 2023. A Formalisation of the Balog–Szemerédi–Gowers Theorem in Isabelle/HOL. In *Proceedings of the 12th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP '23), January 16–17, 2023, Boston, MA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3573105.3575680

## 1 Introduction

The area of formalisation of mathematics with proof assistants has in recent years seen a considerable increase in activity, attracting both computer scientists and mathematicians. This interest is motivated not only by verification purposes, but also by the need to gain new insights on proofs, test the limitations of our available tools, and expand the libraries of formal proofs by enriching them with advanced, research-level mathematics. This thus paves the way for the future creation of tools which – with the promising assistance of AI technology – would support research mathematicians in their creative work.

Significant formalisation work in combinatorics and additive number theory has recently been achieved in numerous proof assistants. The formalisation of the solution to the Cap Set Problem (a 2017 result by Ellenberg and Gijswijt [15]) by Dahmen, Hölzl and Lewis [8] in Lean constitutes one important milestone. Many other profound, albeit less recent, theorems in this area involving arithmetic progressions, Ramsey-type results and extremal graph theory have also been recently formalised. These include: Szemerédi's Regularity Lemma and Roth's Theorem on Arithmetic Progressions formalised by Edmonds, Koutsoukou-Argyraki and Paulson in Isabelle/HOL [11–13] – simultaneously and independently formalised in Lean by Dillies and Mehta [9]; the Hales-Jewett Theorem formalised in Isabelle/HOL by Sulejmani, Eberl and Kreuzer [37], building on a recent formalisation of van der Waerden's Theorem in Isabelle/HOL by Kreuzer and Eberl [22]; the same two results in Lean's mathlib contributed by Wärn by deriving van der Waerden's Theorem as a corollary of the more general Hales-Jewett Theorem. Furthermore, in extremal combinatorics, the Kruskal-Katona theorem has been formalised by Mehta in Lean [23] and the Sunflower Lemma of Erdős and Rado has been formalised by Thiemann in Isabelle/HOL [39].

Specifically, material in additive combinatorics involving the study of the size of sumsets of finite subsets of abelian groups, that is more directly related to the present formalisation, has been recently formalised in Isabelle/HOL. In particular, Koutsoukou-Argyraki and Paulson have formalised the Plünnecke–Ruzsa Inequality [21] and Khovanskii's Theorem [20]. Bakšys and Koutsoukou-Argyraki have formalised Kneser's Theorem, also deriving the Cauchy–Davenport Theorem as a Corollary [2].

The Balog–Szemerédi–Gowers Theorem, whose formalisation we present in this paper, is a profound result in this area and its proof moreover involves tools from extremal graph theory and probability theory. This is the first formalisation of the Balog–Szemerédi–Gowers Theorem in any proof assistant to our knowledge. We give some background information on the theorem below. Before presenting the statement, we start by introducing the following basic definitions.

**Definition 1.1.** Let $A, B$ be finite subsets of an abelian group. The *sumset* $A+B$ is the set $\{a+b \mid a \in A, b \in B\}$. Analogously, the *difference set* $A - B$ is the set $\{a - b \mid a \in A, b \in B\}$.

**Definition 1.2.** Let $G$ be an abelian group. An *additive quadruple* in $G$ is a quadruple $(a, b, c, d) \in G^4$ such that $a + b = c + d$. The *additive energy* of a subset $A$ of $G$ is the number of additive quadruples in $A^4$ divided by $|A|^3$.

A deep result shown by Balog and Szemerédi in 1994 attests that every finite subset $A$ (of given additive energy) in an abelian group must contain a subset $A'$ of $A$ (the cardinality of $A'$ depending on the additive energy of $A$) so that the cardinality of $A'$ is large but the cardinality of the sumset $A' + A'$ is small [6]. In 2001, Gowers gave a new proof of this result with much better bounds on the cardinalities. This constituted a key ingredient of his work on a new proof of the celebrated Szemerédi's Theorem on arithmetic progressions, where he also derived the first effective bounds for Szemerédi's Theorem [16]. The latter bounds provide an estimate on the cardinality of subsets of the naturals that contain no $k$-term arithmetic progressions. Gowers's resultant key improvement of the Balog–Szemerédi Lemma has thus since been known as the Balog–Szemerédi–Gowers Theorem.

The Balog–Szemerédi–Gowers Theorem is today recognised in additive combinatorics for both its proof, which introduced several new incredibly useful techniques, and as a valuable tool in its own right in current research. Tao and Vu's book [38] offer many examples of its early applications to various proofs. Various different expositions – as well as different versions and refinements – of the Balog–Szemerédi–Gowers Theorem are available in the literature, e.g. by Zhao [41], and Sudakov, Szemerédi and Vu [36]. For our formalisation, we followed a proof presented in the 2022 lecture notes by Gowers "Introduction to Additive Combinatorics" for Part III of the Mathematics Tripos taught at the University of Cambridge, which are freely available online [17]. The main statement as we formalised it (corresponding to Corollary 2.19 in the aforementioned notes) reads:

**Theorem 1.3.** *Let $A$ be a finite subset of an abelian group. Suppose that $A$ has additive energy $2c$ for some $c > 0$. Then $A$ has a subset $A'$ so that $|A'| \geq c^2|A|/4$ and $|A' - A'| \leq 2^{30}|A|/c^{34}$.*

We moreover show and formalise an analogous version of the above result for the cardinality of the sumset $|A' + A'|$ instead of the difference set $|A' - A'|$.

The proof assistant of our choice for this formalisation is Isabelle/HOL [24, 29, 30], an interactive theorem prover encoding higher-order logic which features the formal proof language Isar [40] that admits structured proofs. It supports powerful automation through the Sledgehammer proof automation interface [31], as well as counterexample-finding tools. Locales, Isabelle's module system [3], are ideal for formalising mathematical objects and contexts, providing persistent contexts consisting of parameters and assumptions, which can be extended on, combined, and indirectly inherited. Built on the Isabelle libraries, the Archive of Formal Proofs (AFP) contains an extensive collection of formalised material in mathematics, computer science, and logic. As of November 30 2022, the AFP contains 714 entries corresponding to over 3.668.600 lines of code in the above areas combined.

The interplay between different mathematical areas in the proof of the Balog–Szemerédi–Gowers Theorem, namely applications of graph theory and probability theory to additive combinatorics, is of profound mathematical interest. At the same time this makes the formalisation process more interesting, rich and technically challenging. As such, in addition to the formalisation of the Balog–Szemerédi–Gowers Theorem – an essential tool for formal combinatorics libraries in its own right – this work also presents a number of other contributions. This includes additions to libraries on additive combinatorics, probability theory and significant work on graph theory in Isabelle/HOL. In particular, as we will discuss in detail in Section 2, a new undirected graph theory library which is both flexible and extensible was developed by Edmonds [10] with the incentive of this work. Furthermore, we provide a number of interesting contributions to the broader context of the formalisation process for mathematics. Notably, this includes the development of probabilistic methods in Isabelle/HOL – more specifically the *Dependent Random Selection Method* that will be employed in the proof – and a case study on the use of locales, Isabelle's module system, to manage the interplay between different mathematical contexts.

Our formal proof development can be found on the Archive of Formal Proofs [19]. This paper gives an outline of our formalisation and is organised as follows: in Section 2, we discuss the new background material that we formalised for the needs of the main proof; in Section 3, we present the main ideas of Gowers's proof of the Balog–Szemerédi–Gowers Theorem commenting on our formalisation in parallel; in Section 4, we present the final argument for the completion of the proof of the main statement as well as the analogous statement on sumsets; in Section 5, we provide some supplementary, related results of interest that we formalised and which were not required for the main proof; finally, in Section 6, we discuss the broader contributions of this work from the perspective of the formalisation process, with a specific focus on locales to treat the interplay between mathematical contexts, before concluding in Section 7.

## 2 New Background Material Formalised

As we have seen, the main statement that we formalise concerns sumsets of finite subsets of an abelian group. Our development builds on the first basic development of sumset theory recently formalised by Koutsoukou-Argyraki and Paulson [21], which, in turn, builds on Ballarin's algebra development [4].

The proof, however, also makes use of graph-theoretic and probabilistic arguments, in addition to further required background material in additive combinatorics. We therefore had to develop formalisations of a considerable amount of new background material in these areas, which this section describes.

### 2.1 A Triangle Inequality for Sumsets

As it can be seen in our code [19], the assumption that $G$ is an abelian group throughout our formalisation is implemented within the locale *additive_abelian_group*. In particular, we fix $G$ to be the abelian group with chosen symbols for addition and zero. Here we present the Isabelle code that defines this locale:

```
locale additive-abelian-group = abelian-group G (⊕) 0
  for G and addition (infixl ⊕ 65)  and zero (0)
```

The main proof of the Balog–Szemerédi–Gowers Theorem as presented in Gowers's notes [17] made use of the Ruzsa triangle inequality [17, 32, 33] (already formalised by Koutsoukou-Argyraki and Paulson [21]), which attests the following:

**Lemma 2.1.** *Let $A, B, C$ be finite subsets of an abelian group $G$. Then,*

$$|A|\ |B-C| \leq |A-B|\ |A-C|.$$

As we will explain in more detail in Section 4 and the end of Section 5, for an alternative version of the main statement for sumsets instead of difference sets, we made use of the following triangle inequality for sumsets that does *not* directly follow from the Ruzsa triangle inequality. We present it together with its formalisation:

**Lemma 2.2.** *Let $A, B, C$ be finite subsets of an abelian group $G$. Then,*

$$|A|\ |B+C| \leq |A+B|\ |A+C|.$$

```
lemma triangle-ineq-sumsets:
  assumes finite A and finite B and finite C and A ⊆ G and
  B ⊆ G and C ⊆ G
  shows card A ∗ card (sumset B C) ≤
  card (sumset A B) ∗ card (sumset A C)
```

For its formalisation, we followed a proof presented in Zhao's book [41] (see Remark 7.2.2, Corollary 7.3.6 [41]). As in Zhao's book, we deduced the above lemma using a simple application of the following lemma (corresponding to Lemma 7.3.4 [41]), which we present alongside its formalisation:

**Lemma 2.3.** *Let $X$ and $B$ be finite subsets of an abelian group $G$ with $X \neq \emptyset$. Suppose that $\frac{|Y+B|}{|Y|} \geq \frac{|X+B|}{|X|}$ for all nonempty subsets $Y \subseteq X$. Then, for any nonempty finite subset $C$ of the abelian group,*

$$\frac{|X+C+B|}{|X+C|} \leq \frac{|X+B|}{|X|}.$$

```
lemma triangle-ineq-sumsets-aux:
  fixes X B Y :: 'a set
  assumes finite X and finite B and X ⊆ G and B ⊆ G and
  X ≠ {} and ⋀ Y. Y ⊆ X ⟹ Y ≠ {} ⟹
  card (sumset Y B) / card Y ≥  card (sumset X B) / card X and
  finite C and C ≠ {} and C ⊆ G
  shows card (sumset X (sumset C B)) / card (sumset X C) ≤
  card (sumset X B) / card X
```

The argument to prove the above lemma involves an induction on the cardinality of $C$, which we were able to naturally translate into Isabelle/HOL. Nevertheless, we found that on a few occasions, we had to treat the case of the empty set separately, which was omitted from our source. The formal proof spans around 240 lines (versus around 32 in the book exposition), hence its de Bruijn factor can be estimated at around 7.5.

### 2.2 More Material in Additive Combinatorics

A considerable amount of useful technical results in additive combinatorics were formalised, such as a number of basic facts and technical lemmas on the notions of additive quadruples and additive energy and their properties. Here we present the formalised definitions (recall Definition 1.2)

```
definition additive-quadruple:: 'a ⇒'a ⇒'a ⇒ 'a ⇒ bool where
  additive-quadruple a b c d ≡ a ∈ G ∧ b ∈ G ∧ c ∈ G ∧ d ∈ G ∧
  a ⊕ b = c ⊕ d
definition additive-energy:: 'a set ⇒ real where
  additive-energy A ≡ card (additive-quadruple-set A) / (card A)^3
```

The intuition behind the definition of the additive energy is that the cube of the cardinality of the set is a trivial upper bound on the number of additive quadruples in the set, because for every triplet of elements $(a, b, c)$ in the set, there is clearly at most one element $d$ in the set such that $(a, b, c, d)$ will be an additive quadruple. Indeed:

**lemma** *additive-energy-upper-bound*: *additive-energy A ≤ 1*

Much of the new material introduced within the area of additive combinatorics in this development was rather proof-specific, i.e. catered to the needs of the main proof of the Balog–Szemerédi–Gowers Theorem. In particular, we will make use of the following auxiliary function $f_{diff}$ as well as the notion of a $\theta$-popular difference. These definitions, along with their formalisations, are given below:

**Definition 2.4.** Let $A$ be a finite subset of an abelian group $G$. For each $d \in G$ define $f_{diff}(d)$ to be the number of pairs $(a, b) \in A \times A$ such that $a - b = d$. We say that $d \in G$ is a $\theta$-popular difference if $f_{diff}(d) \geq \theta|A|$.

**definition** *f-diff* :: *'a ⇒'a set ⇒ nat* **where**
*f-diff d A ≡ card* {(*a*, *b*) | *a b*. *a* ∈ *A* ∧ *b* ∈ *A* ∧ *a* ⊖ *b* = *d*}
**definition** *popular-diff* :: *'a ⇒ real ⇒ 'a set ⇒ bool* **where**
*popular-diff d ϑ A ≡ f-diff d A ≥ ϑ ∗ of-real* (*card A*)

Analogous such notions for sums were introduced as well. We moreover include a considerable number of lemmas on various properties of these objects.

### 2.3 A New Graph Theory Library

As we will see in the sketch of the proof in Section 3, a key argument of the proof uses a graph-theoretic auxiliary construct. In particular, a bipartite graph fulfilling certain properties is defined, so that each one of its parts is a copy of a finite subset of an abelian group. This means that the vertices of each part are seen as the elements of the finite subset of the abelian group.

There are multiple existing formalisations of graph theory in the Isabelle AFP. Of note, this includes Noschinski's undirected graph theory basics in the Girth Chromatic AFP development [26], a general purpose directed graph theory library [28], and formalisations specific to various algorithms, such as Dijkstra's [25]. However, none of these are suitable for this project. Formalisations of directed graph theory increase the complexity of proofs in undirected graph theory due to the use of more complex Isabelle structures (such as records). Alternatively, the existing undirected graph theory formalisations form part of specific theory developments, and as such are limited in their definitions. For example, formalisations such as Noschinski's undirected graph basics for the Girth Chromatic Theorem [26], have successfully been built on to formalise notable results in graph theory – such as Szemerédi's Regularity Lemma [12, 13]. However, it has a foundational restriction as vertices are defined as type synonyms of natural numbers. A key argument of this work requires an auxiliary graph-theoretic construct where vertices must be seen as elements of a generic abelian group. Hence, it is clear that a formalisation with a notable restriction on vertex type cannot be used in this context.

To overcome this type constraint and treat the graph-theoretic aspects of the proof, we thus employ a new undirected graph theory library developed by the third author with this incentive [10]. The new graph theory library does not impose any type restrictions on the vertices, thus providing the flexibility with respect to vertex type that we require. It further aims to maintain the simplicity of the set-based representation of undirected edges, with many definitions inspired by those in Noschinski's development [26], while being much more flexible and extensible than past smaller formalisations. This is done by utilising a locale-centric approach, similar to previous work by Ballarin [5], and Edmonds and Paulson [14]. By using this approach, we were able to model multiple types of graphs as locales, including those necessary for the Balog–Szemerédi–Gowers development.

The library [10] includes many different core graph theory definitions, beyond those needed for this specific development [19], with the aim that it could serve in the future as a general purpose library for undirected graph theory developments. In total, it consists of approximately 2600 lines of code. We provide further detail on some of the specifics relevant to this development in the remainder of this section.

#### 2.3.1 On Graphs with Loops.

A notable observation when first examining the use of graph theory in the proof of the Balog–Szemerédi–Gowers Theorem is the use of an undirected graph with loops. This provided a further motivation for the development of a more flexible graph theory library, as previous work [26] strictly modelled an edge as a set of size two, which clearly cannot represent loops.

We first introduce a basic graph system locale, which does not yet restrict the edge size, but simply introduces the well-formed assumptions on edges. Edges are modelled as subsets of the vertex set, hence the *'a edge* type is simply a type synonym for a set of elements of a generic fixed type.

**locale** *graph-system* =
**fixes** *vertices* :: *'a set* (*V*)
**fixes** *edges* :: *'a edge set* (*E*)
**assumes** *wellformed*: *e* ∈ *E* ⟹ *e* ⊆ *V*

When using the primary set-based representation of graph edges, we model loops as singleton sets. Hence, the locale now restricts the graph edges to size one or two.

**locale** *ulgraph* = *graph-system* +
**assumes** *edge-size*: *e* ∈ *E* ⟹ *card e* > *0* ∧ *card e* ≤ *2*

This approach enables us to define many basic properties of a graph with no particular adjustments, so that they can also be used in a simple graph setting. Hence the majority of basic properties were defined in this locale context. This includes anything from neighbourhood to connecting paths to edge density (i.e. the number of edges between two vertex subsets divided by the product of their cardinalities). The one exception of this is degree, which intuitively refers to the number of edge ends attached to a vertex. In a simple graph context, this simply means the number of vertices a vertex is incident to. For classic lemmas such as the Handshake Lemma to be maintained, each loop must contribute two to the degree count, hence a modification was required (in the following, *sedges* refers to simple edges, that is, edges with no loops):

```
definition degree :: 'a ⇒ nat where
  degree v ≡ card (incident-sedges v) + 2 ∗ (card (incident-loops v))
lemma degree-no-loops[simp]: ¬ has-loop v ⟹
  degree v = card (incident-edges v)
```

A simple graph also directly builds the graph system locale, with the more constrained parameter restricting the cardinality of the edge set to two. It is simple to show this satisfies the more general assumption of a graph with loops:

```
sublocale sgraph ⊆ ulgraph V E
 by (unfold-locales)(simp add: two-edges)
```

By proving this relation indirectly, we simplify the assumptions the simple graph locale carries around, while ensuring it inherits all of the definitions and lemmas previously defined for the more general class of graphs with loops.

### 2.3.2 Bipartite Graphs.

A bipartite graph is defined by the following locale, using the existing *all-bi-edges* definition which defines the set of all possible edges between two vertex sets:

```
locale bipartite-graph = graph-system +
  fixes X Y :: 'a set
  assumes partition: partition-on V {X, Y}
  assumes ne: X ≠ Y
  assumes edge-betw: e ∈ E ⟹ e ∈ all-bi-edges X Y
```

We choose to model the graph using two explicit parameters for the partition of the vertex set. This generally simplified both the definition and theorem statements so that they better reflected the pen-and-paper source text, however it meant that statements could not be easily generalised as to apply to either set. We were able to establish an easy pattern for symmetric reasoning on bipartite graph properties to overcome this, including for "without loss of generality" statements, by using locale interpretations. These interpretations made use of the basic fact that swapping the vertex sets will still result in a bipartite graph instance:

```
lemma bipartite-sym: bipartite-graph V E Y X
```

A bipartite graph can be shown to be a simple graph indirectly, again ensuring that it inherits all the earlier defined properties of simple graphs and graphs with loops:

```
sublocale bipartite-graph ⊆ sgraph
```

Gowers [17] defines a number of concepts on bipartite graphs. The *codegree* $d(x, x')$ of two vertices is defined to be the number of vertices joined to both $x$ and $x'$, and the *normalised codegree* $\delta(x, x')$ given two vertices in $X$ is their codegree multiplied by $|Y|^{-1}$. By definition, for a bipartite graph, these definitions will be equal to 0 if given two vertices from different partition sets. Note that Gowers [17] uses $\delta$ singularly to represent density, whereas $\delta(x, y)$ represents the normalised codegree. We use the same notation to remain consistent with the source material. The codegree definition is straightforward to formalise (note as we are doing this in a bipartite, and therefore simple, graph environment, there is no need to consider loops):

```
definition codegree:: 'a ⇒ 'a ⇒ nat where
  codegree v u ≡ card {x ∈ V . vert-adj v x ∧ vert-adj u x}
```

The normalised codegree requires slightly more work, as it has two analogous definitions depending on whether the vertices are in $Y$ or $X$. As such, we first define a general definition given a set $S$ and then we are able to show the form of the definition specific to $X$ or $Y$.

```
definition codegree-normalized:: 'a ⇒ 'a ⇒ 'a set ⇒ real where
  codegree-normalized v u S ≡ codegree v u / card S
```

## 2.4 Probability Theory

Probability theory is well-developed in the main Isabelle libraries, with a vast amount of measure theory formalised that in turn is used by many existing AFP entries. This includes two specific entries in combinatorics, which, to our knowledge, are the only existing formalisations of the probabilistic method in combinatorics currently available in any proof assistant. Noschinski proved the classic Girth Chromatic Theorem [27], followed by a formalisation of the Random Graph Subgraph Threshold Theorem for graph properties by Hupel [18].

In the Girth Chromatic development [27], Noschinski developed a locale-based theory on edge spaces, however this proved unnecessary in the context of the Balog–Szemerédi–Gowers development. Here, we only require a much simpler probability measure: a uniform count measure over the vertex set of a graph. There are some basic probability facts in the Random Graphs development [18] which we use in this context.

For this development, we formalise a number of further auxiliary facts on probability theory which are applicable generally. These facts are all in the context of the *prob-space* locale, within which $M$ is a parameter representing the measure of the probability space. It further specifies a number of useful abbreviations within the locale context, such as expectation, which is defined as the Lebesgue integral over the measure $M$.

Firstly, the probabilistic method in combinatorics enables us to use inequalities on the expectation of a function, to show that there must exist a concrete object for which that inequality holds. As such, we prove an "obtains" lemma to enable easy formal reasoning on such statements in a proof:

```
lemma expectation-obtains-ge:
  fixes f :: 'a ⇒ real
  assumes M = uniform-count-measure X and finite X
  assumes expectation f ≥ c
  obtains x where x ∈ X and f x ≥ c
```

Lemma 3.1 requires a variation on the Cauchy–Schwarz inequality detailed in Gowers's notes [17]: given a random variable $X$, for the expected value we have $\mathbb{E}X^2 \geq (\mathbb{E}X)^2$.

We first show this variation using assumptions regarding integrable measures. However, in our main formalisation we only require the simpler context of uniform count measures. A simple lemma attests that a uniform count measure is always integrable on a finite space. As it can be seen in the Isar proof below, this can be used to discharge integrability assumptions before applying the original *cauchy-schwarz-ineq-var* lemma to prove a simpler lemma statement.

```
lemma cauchy-schwarz-ineq-var-uniform:
  fixes X :: 'a ⇒ real
  assumes M = uniform-count-measure S
  assumes finite S
  shows expectation (λ x. (X x)^2) ≥ (expectation (λ x . (X x)))^2
 proof –
  have borel: X ∈ borel-measurable M using assms by simp
  have integrable M X using assms
   by (simp add: integrable-uniform-count-measure-finite)
  then have integrable M (λ x. (X x)^2) using assms
   by (simp add: integrable-uniform-count-measure-finite)
  thus ?thesis using cauchy-schwarz-ineq-var borel by simp
 qed
```

Lastly, we formalise a number of basic facts on expectation over discrete random variable distributions. In many cases, a version of these facts was already available in the measure theory library. However these again had complex assumptions on integrable functions and Bochner integrals, which made the main proofs in the discrete context unnecessarily complex. As such, we use these general facts to formalise variations specifically for uniform count measures. Similar to the Cauchy–Schwarz uniform variation above, this effectively hides unnecessary assumptions and complex notation, which both significantly simplifies later proofs and presents the lemmas in a more recognisable form for a combinatorial setting. One such example is given below:

```
lemma expectation-uniform-count:
  assumes M = uniform-count-measure X and finite X
  shows expectation f = (∑ x ∈ X. f x) / card X
```

## 3 Towards the Proof: Sketch of Gowers's Main Argument

After having introduced the preliminary definitions, we are now ready to sketch the main ideas of the proof and how these interplay in the formalisation.

As we have mentioned before, the proof will rely on an ingenious "detour" via graph theory. Other resources discussing the Balog–Szemerédi–Gowers Theorem, such as Zhao's book [41], sometimes even present the last lemma of this section, Lemma 3.6, as the "graphical" version of the theorem. Key to the proof of this lemma is an auxiliary bipartite graph construct based on the structure of the group, which enables us to translate the graph-theoretic results to group theory.

Following Gowers's exposition, we will start with the graph-theoretic parts of the proof, Lemma 3.1 and Lemma 3.2, which utilise probabilistic arguments. Lemma 3.2 does this via separate technical probabilistic lemmas this section also presents, Lemma 3.3 and Lemma 3.4. The reader will then be able to see the connection of these graph-theoretic arguments with the structure of abelian groups as we proceed with the later parts of the proof sketch. This includes a purely group-theoretic property, Lemma 3.5, before concluding with the key lemma mentioned above, Lemma 3.6.

We start with the following graph-theoretic lemma on the existence of a subset with certain properties in suitably dense bipartite graphs. Its proof utilises the *Dependent Random Selection Method*, a powerful example of the Probabilistic Method [1] in combinatorics. We begin by introducing the lemma and some relevant definitions, before exploring this technique in more detail.

**Lemma 3.1.** *(Lemma 2.13 [17]) Let $H$ be a bipartite graph with finite vertex sets $X$, $Y$ and density $\delta$. Then, for every $c > 0$ there exists a subset $X'$ of $X$ such that $|X'| \geq \delta\,|X|/\sqrt{2}$ and the proportion of pairs $(x, x') \in X' \times X'$ such that for the normalised codegree we have $\delta(x, x') < c$ is at most $2c\delta^{-2}$.*

For our Isabelle formalisation, as is also done within Gowers's original proof, we defined the notion of a *bad pair* to characterise pairs of vertices with the chosen restriction on their normalised codegree as described in the above lemma. Furthermore, *bad-pair-set* is defined as the set of all bad pairs in a vertex set. By formalising the definition of a bad pair outside the proof context, we were able to further prove a number of basic facts to make the formalisation more modular. Additionally, note that *density* is just an abbreviation for *edge-density* introduced in the Bipartite Graphs theory. Here *density* and *edge-density* are used interchangeably as in a bipartite graph the number of edges between the two vertex subsets coincides with the total number of edges of the graph. The Isabelle formalisation of the above lemma thus reads:

```
lemma (in fin-bipartite-graph)
  proportion-bad-pairs-subset-bipartite:
  fixes c::real
  assumes c > 0
  obtains X' where X' ⊆ X and
  card X' ≥ density ∗ card X / sqrt 2 and
  card (bad-pair-set X' Y c) / (card X')^2 ≤ 2 ∗ c / density^2
```

Gowers describes the *Dependent Random Selection Method* as follows [17]: suppose we have a set with certain properties and we want to find a subset with better properties. Choosing that subset purely at random will not necessarily be helpful. Instead, we may be able to find a different distribution on the subsets that is both linked to the structure of the original set and favours the desired improved properties. In this case, instead of picking a purely random subset of $X$ which would not be useful to do, we want to pick some $y \in Y$ at random first. We then let $X'$ be a subset which is the neighbourhood of $y$, and proceed to show it has the desired properties. The formalisation closely mirrors the proof, however requires more specific details. For example, to "pick some $y \in Y$ at random", we must first establish a probability space $P$, which we do via interpretation of the *prob-space* locale using a *uniform-count-measure*. Rather than explicitly picking a $y$, we instead show the required fact on the expectation of the size of $X'$ directly, for which we are able to utilise the earlier formalised variation on the Cauchy–Schwarz inequality and the uniform count measure lemmas on expectation. The setup process and method for stating facts on expectation provide an example of a template for applying probabilistic techniques in Isabelle/HOL.

The second half of the proof deals with some complex inequalities, for which the formal proof required many more intermediate steps than the original source. Notably, the inequality on the expected number of bad pairs in $X'$ is presented in a single sentence in Gowers's notes [17], intuitively making use of facts on the linearity of expectation. In order to apply these facts formally, we represented the cardinality of a set as the summation over the indicator functions of elements of a superset, and were then able to reason formally using multiple applications of linearity. The existing formalisation of *indicator* functions in Isabelle proved valuable to this end, as a number of probabilistic methods on indicator functions had previously been formalised.

A helpful observation, which we utilise later in our formalisation, is that the codegree $d(x, x')$ equals the number of paths of length two between the vertices $x$ and $x'$:

```
lemma codegree-is-path-length-two:
  codegree x x′ = card {p . connecting-path x x′ p ∧
  walk-length p = 2}
```

By this observation, Lemma 3.1 can be interpreted to guarantee that, given a dense bipartite graph, we can restrict one of its vertex sets to a large subset in which *almost all* pairs of vertices are joined by many paths of length two. Lemma 3.1 together with the above lemma is not yet strong enough for our purposes. We will, however, use it to prove a similar statement which guarantees the existence of a large number of walks of length three. It is important to note the transition from paths to walks here in our development. By the conventions of our undirected graph theory library, which we introduced in Section 2.3, a walk on a graph is defined as a list of vertices, where consecutive vertices are connected by an edge, whereas a path is defined to be a non-self-intersecting walk or a cycle. In our formalisation source notes, Gowers uses paths throughout all of the proofs without giving a formal definition. We initially set out to prove all of the results for paths, however quickly realised that they only held for our definition of a walk with a few minor exceptions (e.g. *codegree-is-path-length-two* because in our bipartite graph context walks of length two are also paths). With this in mind, we will now seek to prove a lemma which attests the existence of many walks of length three between *all* pairs where, this time, each member of a pair is in a large subset of a different part of the bipartite graph:

**Lemma 3.2.** *(Lemma 2.16 [17]) Let $H$ be a bipartite graph with finite vertex sets $X, Y$ and density $\delta$. Then there are subsets $X' \subseteq X$, $Y' \subseteq Y$ with $|X'| \geq \delta^2\ |X|/16$ and $|Y'| \geq \delta\ |Y|/4$ such that for every $x \in X'$ and $y \in Y'$, the number of walks of length three between $x$ and $y$ in $H$ is at least $\delta^6\ |X|\ |Y|/2^{13}$.*

```
lemma (in fin-bipartite-graph)
  walks-of-length-3-subsets-bipartite:
  obtains X′ and Y′ where X′ ⊆ X and Y′ ⊆ Y and
  card X′ ≥ (edge-density X Y)^2 ∗ card X / 16 and
  card Y′ ≥ edge-density X Y ∗ card Y / 4 and
  ∀ x ∈ X′. ∀ y ∈ Y′.
  card {p. connecting-walk x y p ∧ walk-length p = 3} ≥
  (edge-density X Y)^6 ∗ card X ∗ card Y / 2^13
```

Initially, we found a variation of Gowers's proof of Lemma 3.2 that does not require the observation given as *codegree-is-path-length-two* lemma above but rather given a pair $(x, y) \in X' \times Y'$ explicitly constructs the required walks of length three from $x$ to $y$. From here, a simple unfolding of the definition of the codegree allowed us to finish the proof. It must be noted that the alternative version of the proof was found simply because it was easier to carry out the formalisation using sets rather than walks (which are defined as lists). Ultimately, we decided to rework the proof to more closely follow the source to make it more readable and modular. This utilised the auxiliary *codegree-is-path-length-two* lemma to build up walks of length three by appending an extra element to paths of length two and resulted in a more modular and general formalisation approach.

The proof of Lemma 3.2 is probabilistic in its essence as we will need to employ two technical probabilistic lemmas, Lemma 3.3 and Lemma 3.4 (in addition to Lemma 3.1). These two technical probabilistic lemmas will give us a way to turn probabilistic statements into lower bounds on the cardinality of certain sets, which we will directly apply to obtain large dense subsets of graphs. We present the two technical probabilistic lemmas below along with their formalisations:

**Lemma 3.3.** *(Lemma 2.14 [17]) Let $X$ be a finite set and let $f : X \longrightarrow [0, 1]$. Assume $\mathbb{E}_x\ f(x) \geq \delta$. Then there are at least $\delta|X|/2$ many elements $x \in X$ such that $f(x) \geq \delta/2$.*

```
lemma (in prob-space) expectation-condition-card-1:
  fixes X::′a set  and f::′a ⇒ real and δ::real
  assumes finite X and ∀ x ∈ X. f x ≤ 1 and
  M = uniform-count-measure X and expectation f ≥ δ
  shows card {x ∈ X. (f x ≥ δ / 2)} ≥ δ ∗ card X / 2
```

**Lemma 3.4.** *(Lemma 2.15 [17]) Let $X$ be a finite set and let $f : X \longrightarrow [0, 1]$ be a function with $\mathbb{E}_x\ f(x) \geq 1 - \alpha$ for some $\alpha > 0$. Then for every $\beta > 0$ the number of $x \in X$ such that $f(x) \geq 1 - \beta$ is at least $(1 - \alpha/\beta)|X|$.*

```
lemma (in prob-space) expectation-condition-card-2:
  fixes X::′a set and β::real and α::real and f:: ′a ⇒ real
```

```
assumes finite X and ⋀ x. x ∈ X ⟹ f x ≤ 1 and
β > 0 and α > 0 and expectation f ≥ 1 − α and
M = uniform-count-measure X
shows card {x ∈ X. f x ≥ 1 − β} ≥ (1− α / β) ∗ card X
```

For the formalisation of these probability theory lemmas we used the exact same underlying measure as described in Section 2.4 and we naturally followed the pen-and-paper proof. For both proofs we employed the lemma *expectation-uniform-count*, presented in Section 2.4.

As we have outlined above, for the proof of Lemma 3.2, we needed to combine graph-theoretic results in the form of Lemma 3.1 with probabilistic results in the form of Lemmas 3.3 and 3.4 – all this while simultaneously passing to multiple auxiliary graphs. To do all this in the proof context, we heavily relied on locales. Using them, we were able to seamlessly define certain auxiliary structures and prove their membership to a particular locale, which allowed us to naturally apply all previous theory on these auxiliary constructions. A good example of this is the following interpretation of the *uniform-count-measure* probability space over one side $X$ of the finite bipartite graph:

```
interpret P1: prob-space uniform-count-measure X
  by (simp add: X-not-empty partitions-finite(1)
  prob-space-uniform-count-measure)
```

Following the above interpretation, using Lemma 3.3 we are able to find a large subset $?X1$ of $X$ with large normalised degree. From here, the pen-and-paper proof restricts the given finite bipartite graph to one with vertex sets $?X1$ and $Y$. In order to reason about this induced graph, we once again use an interpretation:

```
interpret H:
fin-bipartite-graph (?X1 ∪ Y) {e ∈ E. e ⊆ (?X1 ∪ Y)} ?X1 Y
```

We note that to distinguish between the finite bipartite graph fixed by the locale context of the main body of the proof, and the induced finite bipartite graph, we have given our interpretation the name $H$. This allows us to distinguish between definitions and lemmas relating to each of the graphs by using dot notation. Having interpreted the aforementioned induced graph, we still need a way to transport information across these structures. A good example of such a transfer is a straightforward claim that neighbourhoods and degree of vertices in $?X1$ are unchanged:

```
have neighborhood-unchanged: ∀ x ∈ ?X1.
  neighbors-ss x Y = H.neighbors-ss x Y
   using neighbors-ss-def H.neighbors-ss-def vert-adj-def
   H.vert-adj-def by auto
then have degree-unchanged: ∀ x ∈ ?X1. degree x = H.degree x
  using H.degree-neighbors-ssX degree-neighbors-ssX by auto
```

Our formal proof required multiple such statements, which transport information across different structures to deduce the final statement – something almost no author would write down in a pen-and-paper proof. Another interesting use case of an interpretation in this proof, which also inspired us to develop our undirected graph theory library as discussed in Section 2.3, is one of a graph with loops:

```
let ?E-loops = mk-edge ‘{(x, x′) | x x′. x ∈ X2 ∧ x′ ∈ X2 ∧
  (H.codegree-normalized x x′ Y) ≥ ?δ ^ 3 / 128}
interpret Γ: ulgraph X2 ?E-loops
```

Here, we first define the edge set denoted by $?E - loops$ on the set $X2$, which we obtained by extracting a large subset of $?X1$ that has a small number of bad pairs with respect to $H$. Noting that for the graph with loops $\Gamma$ the expectation of normalised degree is high, we finally use Lemma 3.4 to obtain a large subset $X'$ of $X2$ whose vertices all have high normalised degree in $\Gamma$. From here, the proof straightforwardly follows the pen-and-paper proof by finding a large subset $Y'$ of $Y$ with all vertices having high normalised degree into $X'$ and constructing the required number of walks of length three by unfolding the definitions of the auxiliary graphs $H$ and $\Gamma$.

It is now time to turn to the group-theoretic aspects of the proof. We proceed with the following lemma[1] (recall Definition 2.4):

**Lemma 3.5.** *(Lemma 2.17 [17]) Let $A$ be a finite subset of an abelian group $G$ and suppose that the additive energy of $A$ is $2c$ for some $c > 0$. Then the number of $c$-popular differences $d \in G$ is at least $c|A|$.*

```
lemma popular-differences-card: fixes A::′a set and c::real
  assumes finite A and A ⊆ G and additive-energy A = 2 ∗ c
  shows card (popular-diff-set c A) ≥ c ∗ card A
```

**Remark**: Note that in Lemmas 3.5 and 3.6 and thus in the final result presented in Section 4, we have set the additive energy to be equal to $2c$ for some $c > 0$, and we are accordingly considering $\theta$-popular differences where $\theta = c$ so that we are consistent with the original proof where $\theta$ is chosen to be the half of the additive energy of $A$. The appearing numerical discrepancy between the bounds in the notes and in our formalisation is thus only artificial.

For the proof of Lemma 3.5, we made use of certain properties of the function $f_{diff}$ (recall the definition in Section 2.2) in relation to the cardinality of $A$ and the additive energy of $A$ (recall Definition 1.2). In particular, among others, we used lemmas *f-diff-le-card* and *f-diff-card-quadruple-set-additive-energy* and *f-diff-card*, which we give below:

```
lemma f-diff-le-card:
  assumes finite A and A ⊆ G
  shows f-diff d A ≤ card A
lemma f-diff-card-quadruple-set-additive-energy:
  assumes A ⊆ G and finite A
  shows (∑ d ∈ differenceset A A. (f-diff d A)^2) =
  additive-energy A ∗ (card A)^3
```

[1] To avoid confusion, we note that there is a typographical error in the statement of Lemma 3.5 in the online notes [17]: the square should be missing.

```
lemma f-diff-card:
  assumes A ⊆ G and hA: finite A
  shows (∑ d ∈ (differenceset A A). f-diff d A) = (card A)^2
```

Note that in the original source, the lemmas *f-diff-card-quadruple-set-additive-energy* and *f-diff-card* correspond to slightly different statements. In particular, they read: for a finite subset $A$ of an abelian group $G$, $\sum_{d\in G} f_{diff}(d) = |A|^2$ and $\sum_{d\in G}(f_{diff}(d))^2$ equals the number of additive quadruples in $A$ (i.e. the additive energy of $A$ multiplied by $|A|^3$: recall Definition 1.2). While the versions of these lemmas found in Gowers's notes are correct mathematically, we could not translate them verbatim into Isabelle/HOL. This is because in the main library, summation over an infinite set is defined to be equal to 0 and the abelian group $G$ is not restricted to be finite. Hence, we had to restrict our indexing set to the difference set, which is finite and simultaneously contains the support of $f_{diff}$. This indexing restriction meant that we had to make repeated use of the fact that the difference set here is finite:

```
lemma finite-differenceset: finite A ⟹ finite B ⟹
  finite (differenceset A B)
```

The following lemma is the key argument of the main proof. Within its proof, the results from probabilistic graph theory previously presented get introduced to the study of difference sets/sumsets which is our ultimate purpose:

**Lemma 3.6.** *(Lemma 2.18 [17]) Let $A$ be a finite subset of an abelian group $G$ and suppose that the additive energy of $A$ is $2c$ for some $c > 0$. Then $A$ has subsets $B$ and $A'$ with $|B| \geq c^4\ |A|/16$ and $|A'| \geq c^2\ |A|/4$ so that $|A' - B| \leq 2^{13}\ |A|/c^{15}$.*

```
lemma obtains-subsets-differenceset-card-bound:
  fixes A::'a set and c::real
  assumes finite A  and c>0  and A ≠ {} and A ⊆ G and
  additive-energy A = 2 ∗ c
  obtains B and A' where B ⊆ A and B ≠ {} and
  card B ≥ c^4 ∗ card A / 16  and A' ⊆ A and A' ≠ {} and
  card A' ≥ c^2 ∗ card A / 4  and
  card (differenceset A' B) ≤ 2^13 ∗ card A / c^15
```

The proof of Lemma 3.6 makes use of the following auxiliary graph-theoretic construct (where $A$ is as above). We define a bipartite graph $H$ in the following fashion: the vertex sets $X$ and $Y$ of $H$ are defined to be copies of $A$, and vertices $a$ and $b$ are connected by an edge if and only if $b - a$ is a $c$-popular difference. To implement this formally, we had to interpret $H$ as a finite bipartite graph within the proof context, so that we could apply Lemma 3.2. This meant that we had to prove that the copies of $A$ are disjoint and partition the vertex set, hence we made the design decision to work with $A \times \{0\}$ and $A \times \{1\}$ as our *copies* of $A$. We present an excerpt of the above construction and interpretation below:

```
  let ?X = A × {0:: nat}
  let ?Y = A × {1:: nat}
  let ?E = mk-edge ‘ {(x, y)| x y. x ∈ ?X ∧ y ∈ ?Y ∧
  (popular-diff (fst y ⊖ fst x) c A)}
  interpret H: fin-bipartite-graph ?X ∪ ?Y ?E ?X ?Y
```

Further, using Lemma 3.5, we deduce that the density of $H$ must thus be at least $c^2$. Applying Lemma 3.2 we can now obtain subsets $B \subseteq A$, $A' \subseteq A$ of cardinalities bounded from below by $c^4|A|/16$ and $c^2|A|/4$, respectively, such that for every $x \in B$ and $y \in A'$, the number of walks of length three between $x$ and $y$ in $H$ is at least $c^{13}|A|^2/2^{13}$. The number of walks of length three translates into a number of choices so that certain differences of elements in $A$ would all be $c$-popular. Specifically, this means that for each difference $a' - b \in A' - B$, we can find $c^{13}|A|^2/2^{13}$ many $(z, w) \in A^2$ such that $z - b$, $z - w$ and $a' - w$ are all $c$-popular. Although this deduction is very short on pen-and-paper, we found that its formal proof was lengthy, which was also the case for other arguments involving transitions across different structures. In our development the above statement reads:

```
have card-ineq1: ⋀ x y. x ∈ ?B ⟹ y ∈ ?C ⟹
  card ({(z, w) | z w. z ∈ A ∧ w ∈ A ∧
  popular-diff (z ⊖ x) c A ∧ popular-diff (z ⊖ w) c A ∧
  popular-diff (y ⊖ w) c A}) ≥ (c^12) ∗ ((card A)^2) / 2^13
```

It is important to note that in the formal environment, we firstly obtained subsets $?B$ and $?C$ of $A \times \{0\}$ and $A \times \{1\}$, respectively, which we had to project down to subsets of $A$ to obtain $B$ and $A'$, respectively. And although this claim occupied only a sentence in the original source, it spanned 113 lines of Isar code in our formalisation.

To finish the proof sketch, observe that for each sextuple $(p, q, r, s, t, u) \in A^6$ such that $p - q = z - b$ and $r - s = z - w$ and $t - u = a' - w$, we can recover $z$ and $w$ uniquely for fixed $a'$ and $b$. From here, by a straightforward pen-and-paper computation one finds at least $c^{15}|A|^5/2^{13}$ such sextuples in $A^6$ for each $d \in A' - B$, which finally gives us the desired bound on the cardinality of $A' - B$. In the formal environment, we note that the $c$-popular condition on certain differences in a connected walk of length three naturally gives rise to many triples of pairs $((p, q), (r, s), (t, u)) \in A^2 \times A^2 \times A^2$ satisfying the above properties. In order to transfer our information into claims for sextuples, we locally defined an auxiliary bijection between $A^6$ and $A^2 \times A^2 \times A^2$ as presented below:

```
define f:: 'a × 'a × 'a × 'a × 'a × 'a ⇒
  ('a × 'a) × ('a × 'a) × ('a × 'a) where
  f ≡ (λ (p, q, r, s, t, u). ((p, q), (r, s), (t, u)))
```

We then proved that $f$ is injective and maps the appropriate sextuples into triplets of pairs with the required properties, which allowed us to deduce the final claim.

# 4 Completing the Proof of the Balog–Szemerédi–Gowers Theorem

The proof of the main result Theorem 1.3 (which is presented as Corollary 2.19 in the notes [17]) is now a matter of the direct application of the Ruzsa triangle inequality, i.e. Lemma 2.1 (available from [21]) to Lemma 3.6. This is achieved simply by observing that the Ruzsa triangle inequality gives $|B|\ |A' - A'| \leq |A' - B|^2$. The bounds for $|B|$, $|A' - B|$ are obtained from Lemma 3.6, i.e. lemma *obtains-subsets-differenceset-card-bound*, which thus gives the desired estimate. The main statement, as we formalised it in Isabelle/HOL, reads (recall the remark in Section 3):

```
theorem Balog-Szemeredi-Gowers: fixes A::'a set and c::real
  assumes afin: finite A and A ≠ {} and c>0
  and additive-energy A = 2 ∗ c and ass: A ⊆ G
  obtains A' where A' ⊆ A and card A' ≥ c^2 ∗ card A / 4 and
  card (differenceset A' A') ≤  2^30 ∗ card A / c^34
 proof–
  obtain B and A' where bss: B ⊆ A and bne: B ≠ {} and
   bge: card B ≥ (c^4) ∗ (card A)/16  and
   a2ss: A' ⊆ A  and a2ge: card A' ≥ (c^2) ∗ (card (A))/4  and
   hcardle: card (differenceset A' B) ≤ 2^13 ∗ card A / c^15
   using assms obtains-subsets-differenceset-card-bound by metis
  have bg0: (card B :: real) > 0 ⟨proof⟩
  have (card  B) ∗ card (differenceset A' A') ≤
   card (differenceset A' B) ∗ card (differenceset A' B)
   using afin a2ss bss infinite-super ass Ruzsa-triangle-ineq1
   card-differenceset-commute subset-trans sumset-commute
    by (smt (verit, best))
  then have card B ∗ card (differenceset A' A') ≤
    (card (differenceset A' B))^2 ⟨proof⟩
  then have (card (differenceset A' A')) ≤
    (card (differenceset A' B))^2/card B ⟨proof⟩
  moreover have (card (differenceset A' B))^2  ≤
    ((2^13) ∗ (1/c^15)∗(card A))^2  using hcardle  by simp
  ultimately have (card (differenceset A' A')) ≤
    ((2^13) ∗ (1/c^15)∗(card A))^2/(card B) ⟨proof⟩
  moreover have (c^4) ∗ (card A)/16 >0 ⟨proof⟩
  moreover have ((2^13) ∗ (1/c^15) ∗ (card A))^2/(card B) =
    ((2^13)∗ (1/c^15)∗(card A))^2 ∗ (1/(card B)) by simp
  moreover have ((2^13)∗(1/c^15)∗(card A))^2∗(1/(card B))≤
    ((2^13)∗(1/c^15)∗(card A))^2/((c^4)∗(card A)/16) ⟨proof⟩
  ultimately have (card (differenceset A' A')) ≤
   ((2^13) ∗ (1/c^15) ∗ (card A))^2/ ((c^4) ∗ (card A)/16)
    by linarith
  then have (card (differenceset A' A')) ≤
   (2^30) ∗ (card A)/(c^34)
    using card-0-eq assms by (simp add: power2-eq-square)
   then show ?thesis using a2ss a2ge that by blast
 qed
```

In the source material [17], the proof of the main theorem is only three lines long. While this is notably longer, from the Isar proof above we can clearly see how the formalisation follows the pen-and-paper proof: the *obtain* statement is where Lemma 3.6 is applied, the second *have* statement is where the Ruzsa triangle inequality, i.e. Lemma 2.1 is applied, and the remaining *have* statements provide calculation-based detail to reach the required inequality. We can further see where the bound from the *obtain* statement, *hcardle*, is employed in one of these steps via *simp*. All other calculation-based steps were proven using standard numeric/algebraic lemmas and tactics, hence the one-line proof details are replaced by ⟨*proof*⟩ in this paper. This is a classic example of a formal proof requiring a much more detailed series of steps, where on pen-and-paper the leap from one form of an inequality to another is seen as obvious.

We moreover formalise a version of the main theorem for sumsets, that is:

```
theorem Balog-Szemeredi-Gowers-sumset: fixes A::'a set and c::real
  assumes afin: finite A and A ≠ {} and c>0 and
  additive-energy A = 2 ∗ c and ass: A ⊆ G
  obtains A' where A' ⊆ A and card A' ≥ c^2 ∗ card A / 4 and
   card (sumset A' A') ≤ 2^30 ∗ card A / c^34
```

The formalisation has a similar structure to the previous theorem, however, instead of using the Ruzsa triangle inequality, i.e. Lemma 2.1, we applied the different triangle inequality for sumsets, i.e. Lemma 2.2 to *obtains-subsets-differenceset-card-bound*, i.e. Lemma 3.6 in an analogous way. In particular, Lemma 2.2 gives $|B|\ |A' + A'| \leq |A' - B|^2$ and the bounds for $|B|$, $|A' - B|$ are again obtained from *obtains-subsets-differenceset-card-bound*, which thus gives the desired estimate.

## 5 Supplementary Results Formalised

In addition to the Balog–Szemerédi–Gowers Theorem (two versions), we also formalised the proof of some related, simple, supplementary results which were not used in the proofs but we find worth mentioning here. In particular, using an appropriate version of the Cauchy–Schwarz inequality:

```
lemma Cauchy-Schwarz-ineq-sum2:
  fixes f g:: 'a ⇒ real and A:: 'a set
  shows (∑  d ∈ A. f d ∗ g d) ≤
   (∑  d ∈ A. (f d)^2) powr (1/2) ∗ (∑  d ∈ A. (g d)^2) powr (1/2)
```

we formalised the proof of a lower bound on additive energy (corresponding to Proposition 2.11 [17]), attesting that for a finite, nonempty subset $A$ of an abelian group with $|A+A| \leq C|A|$ for some real constant $C$, the additive energy of $A$ is at least $1/C$:

```
proposition additive-energy-lower-bound-sumset:
  fixes C::real
  assumes finite A and A ⊆ G and
  (card (sumset A A)) ≤ C ∗ card A and card A ≠ 0
  shows additive-energy A ≥ 1/C
```

We also show that we can reach the same conclusion as above using the alternative assumption $|A - A| \leq C|A|$ (the proof is analogous with the only difference being the use of the function *f-diff* instead of the analogous *f-sum* within the proof).

Finally, we show two more additional lemmas. In particular, we formalise a lemma analogous to *popular-differences-card* (Lemma 3.5), referring to the analogous function *f-sum* instead of *f-diff*:

```
lemma popular-sums-card:
  fixes A::'a set and c::real
  assumes finite A and additive-energy A = 2 ∗ c and A ⊆ G
  shows card (popular-sum-set c A) ≥ c ∗ card A
```

Similarly, we formalise a lemma analogous to the key lemma: *obtains-subsets-differenceset-card-bound* (Lemma 3.6):

```
lemma obtains-subsets-sumset-card-bound:
  fixes A::'a set and c::real
   assumes finite A  and c>0  and A ≠ {} and A ⊆ G and
  additive-energy A = 2 ∗ c
  obtains B and A' where B ⊆ A and B ≠ {} and
  card B ≥ c^4 ∗ card A / 16  and A' ⊆ A and A' ≠ {} and
  card A' ≥ c^2 ∗ card A / 4  and
  card (sumset A' B) ≤ 2^13 ∗ card A / c^15
```

The latter follows from lemma *popular-sums-card* in the same way lemma *obtains-subsets-differenceset-card-bound* follows from lemma *popular-differences-card* and by substituting *f-diff* with the analogous function for sums *f-sum* , popular differences with the analogous notion for popular sums, and interchanging ⊕ with ⊖. The interest here lies in the observation of the duality between sumsets and difference sets. Exploiting this duality was actually facilitated by the formal proofs, as making the above described substitutions in the formal proofs of *popular-differences-card* and *obtains-subsets-differenceset-card-bound* to obtain the analogous versions for sumsets, i.e. *popular-sums-card* and *obtains-subsets-sumset-card-bound* respectively was straightforward. This method, however, could not be directly applied to obtain the sumset version of the main result, as already explained in Section 4. Rather, an application of the different triangle inequality in the proof, i.e. Lemma 2.2 was necessary for the sumset version due to the parity behaviour (please see Comment 7.2.2 [41]).

Note that the difference set version of the main result theorem *Balog-Szemeredi-Gowers* (i.e. Theorem 1.3) could also have been alternatively shown with the sumset version of Lemma 3.6 above, i.e. with *obtains-subsets-sumset-card-bound*. This would give bounds for $|B|$, $|A' + B|$ via the inequality $|B|\ |A' - A'| \leq |A' + B|^2$ which would again follow from the Ruzsa triangle inequality, i.e. Lemma 2.1. In fact, the difference set version of the main theorem Theorem 1.3 could be shown with either version of Lemma 3.6 (for sumsets or difference sets) but would require the Ruzsa triangle inequality (Lemma 2.1), while the sumset version of the main theorem Theorem 1.3 could also be shown with either version of Lemma 3.6 (for sumsets or difference sets) but would require the alternative triangle inequality for sumsets (Lemma 2.2).

## 6 Discussion

The formalisation process for the Balog–Szemerédi–Gowers Theorem presented some interesting technical challenges, however, thanks to the proof assistant language and tools, it remained relatively straightforward to complete. This is notable for a result of such mathematical significance. In this section, we discuss some of the key contributions of the formalisation beyond the successful formal proof of the final theorem statement.

As discussed throughout Section 2 and 3, we handled the interplay between the different mathematical areas that contributed different elements of the proof, namely graph theory (also considering graphs with loops), probability theory and additive combinatorics, by implementing an appropriate use of locales, which provide a practical module system. Firstly, locales were central to the formalisation of the new undirected graph theory library presented in Section 2.3. This library provides another example of a locale-centric formalisation of a mathematical hierarchy, along the lines of earlier work in both combinatorics and algebra [5, 7, 14]. As noted in prior work, this approach again proved to be both flexible and extensible, while also enabling natural mathematical notation of different structures. Notably, the bulk of this graph theory library [10] was completed in under two weeks and was easily integrated into the existing sketch of the formalisation of the Balog–Szemerédi–Gowers Theorem which was already in progress. This further demonstrates the ease of working with locales to both build such a library, and express and prove various statements in an applied theorem. Secondly, locales continued to be essential in managing the interplay of different mathematical areas. Throughout the proof sketch, we were simultaneously working with locales that represented additive abelian groups, different types of graphs and a probability space. Typically, usage of locales previously has focused on working with one locale at a time, either in its context or on a single instance of a locale. In this formalisation, we have demonstrated how we can use multiple different interpretations of different locale structures to transport results across different contexts, with key examples detailed in Section 3 in the proofs of both Lemma 3.2 and Lemma 3.6. This highlights the power of local interpretations inside proof contexts when using locales, as well as how locale-specific definitions and theorems can be easily used outside of their locale context, further supporting locales as the structure of choice when defining mathematical objects and hierarchies.

Notably, as detailed in Section 3, this formalisation also provides several examples of the application of the probabilistic method to combinatorics in a formal environment, which has only been explored in two other earlier formalisations to the best of our knowledge [18, 27]. We highlight how interpreting (potentially many) instances of the *prob-space* locale, enable us to set up a proof such that we can utilise the powerful previous developments on measure and probability theory in Isabelle's libraries. In combinatorics, the choice and setup of a probability space is usually implicit in a pen-and-paper text, and as such we aim to provide an example of the basic set up that the formal environment requires to follow the same such reasoning.

Isabelle's advanced automation significantly assisted us during the formalisation process; indeed, as the reader would easily notice by inspecting our theory file, there are many instances of Sledgehammer-generated proofs, e.g. proofs by *metis* or by *smt*. Additionally, the Isar proof language proved to be very user-friendly during the formalisation process, enabling us to structure the proof in an easily readable and accessible way. This was particularly important given that there were three contributors to the project with different backgrounds. Moreover, as seen in the formal proof provided in Section 4, it also enabled us to complete many parts of the formalisation in line with the original pen-and-paper proof. It is also worth mentioning that, especially in the first stages of the project, we made frequent use of the search engine for the Isabelle libraries and AFP, SErAPIS[2] [34, 35], which proved to be of valuable help.

Keeping the new graph theory library [10] separate[3], the full proof of the Balog–Szemerédi–Gowers Theorem spanned around 1900 lines of Isar code. This line estimate includes all necessary preliminaries developed e.g. on graphs with loops, bipartite graphs, probability space theory, various simple technical lemmas, definitions and elementary material in additive combinatorics, and excludes the supplementary material and results shown that were not required for the main proof of the Balog–Szemerédi–Gowers Theorem such as these mentioned in Section 5, the alternative version of the main theorem (for sumsets) and the proof of the triangle inequality for sumsets, i.e. Lemma 2.2 that was used for it. Considering around 137 lines (around 4 pages) of mathematical text in Gowers's notes [17] restricted to covering the proof of the Balog–Szemerédi–Gowers Theorem, the de Bruijn factor of our formalisation can thus be estimated at around 13.9.

Notably, the project to formalise the proof of the Balog–Szemerédi–Gowers Theorem was completed in less than two months in total (including the separate graph theory library). All three authors worked on the formalisation, including equal contributions from the second author who was completely new to Isabelle at the initial stage. Hence this two-month period also includes the time it took the second author to become familiar with Isabelle.

## 7 Concluding Comments

We have described our formalisation in Isabelle/HOL of the Balog–Szemerédi–Gowers Theorem, a profound result in additive combinatorics with many significant applications – most notably in an effective version of the celebrated Szemerédi's Theorem. Our formalisation, the first of this result in any proof assistant to our knowledge, moreover motivated the development of useful formalisations of essential background material in graph theory, probability theory and additive combinatorics that could become useful in relevant future developments. The successful completion of this work is an indication that Isabelle/HOL has the capacity for formalising modern and advanced mathematical material involving a combination of different mathematical areas. Such an interplay, as we explained, can be implemented very efficiently by the use of Isabelle's locales in a modular and flexible way. We have moreover demonstrated that this can be achieved within a reasonable time span and even by authors without much prior experience: Isabelle's ecosystem, which includes its advanced automation, the Isar formal proof language as well as efficient search features, makes formalisation of mathematics smooth and accessible. At the same time, the Isabelle libraries and the AFP currently offer a rich, robust collection of formalised material that we can build on, so the time is ripe for the formalisation of more advanced mathematics.

## Acknowledgements

Angeliki Koutsoukou-Argyraki is funded by the ERC Advanced Grant ALEXANDRIA (Project GA 742178) funded by the European Research Council and led by Lawrence C. Paulson (University of Cambridge, Department of Computer Science and Technology). Mantas Bakšys received funding for his internship supervised by Koutsoukou-Argyraki by the Cambridge Mathematics Placements (CMP) Programme and by the ALEXANDRIA Project. Chelsea Edmonds is jointly funded by the Cambridge Trust (Cambridge Australia Scholarship) and a Cambridge Department of Computer Science and Technology Premium Research Studentship. We thank the anonymous referees for their useful comments that significantly improved the presentation of the paper.

## References

[1] Noga Alon and Joel H. Spencer. 2016. *The Probabilistic Method* (4th ed.). Wiley, Hoboken, N.J.

[2] Mantas Bakšys and Angeliki Koutsoukou-Argyraki. 2022. Kneser's Theorem and the Cauchy–Davenport Theorem. *Archive of Formal Proofs* (November 2022). https://isa-afp.org/entries/Kneser_Cauchy_Davenport.html, Formal proof development.

[3] Clemens Ballarin. 2010. Tutorial to Locales and Locale Interpretation. In *Contribuciones Científicas en Honor de Mirian Andrés Gómez*. University of Rioja, 123–140. Online at https://dialnet.unirioja.es/servlet/articulo?codigo=3216664.

[4] Clemens Ballarin. 2019. A Case Study in Basic Algebra. *Archive of Formal Proofs* (August 2019). https://isa-afp.org/entries/Jacobson_Basic_Algebra.html, Formal proof development.

[5] Clemens Ballarin. 2020. Exploring the Structure of an Algebra Text with Locales. *Journal of Automated Reasoning* 64, 6 (August 2020), 1093–1121. https://doi.org/10.1007/s10817-019-09537-9

[6] Antal Balog and Endre Szemerédi. 1994. A statistical theorem of set addition. *Combinatorica* 14 (1994), 263–268. https://doi.org/10.1007/BF01212974

[7] Anthony Bordg, Lawrence C. Paulson, and Wenda Li. 2022. Simple Type Theory is not too Simple: Grothendieck's Schemes Without Dependent Types. *Experimental Mathematics* 31, 2 (2022), 364–382. https://doi.org/10.1080/10586458.2022.2062073

[8] Sander R. Dahmen, Johannes Hölzl, and Robert Y. Lewis. 2019. Formalizing the Solution to the Cap Set Problem. In *10th International*

[2] https://behemoth.cl.cam.ac.uk/search/

[3] but including the special graph-theoretic prerequisites introduced in the Balog–Szemerédi–Gowers Theorem development itself

*Conference on Interactive Theorem Proving (ITP 2019) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 141)*, John Harrison, John O'Leary, and Andrew Tolmach (Eds.). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 15:1–15:19. https://doi.org/10.4230/LIPIcs.ITP.2019.15

[9] Yaël Dillies and Bhavik Mehta. 2022. Formalising Szemerédi's Regularity Lemma in Lean. In *13th International Conference on Interactive Theorem Proving (ITP 2022) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 237)*, June Andronick and Leonardo de Moura (Eds.). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 9:1–9:19. https://doi.org/10.4230/LIPIcs.ITP.2022.9

[10] Chelsea Edmonds. 2022. Undirected Graph Theory. *Archive of Formal Proofs* (September 2022). https://isa-afp.org/entries/Undirected_Graph_Theory.html, Formal proof development.

[11] Chelsea Edmonds, Angeliki Koutsoukou-Argyraki, and Lawrence C. Paulson. 2021. Roth's Theorem on Arithmetic Progressions. *Archive of Formal Proofs* (December 2021). https://isa-afp.org/entries/Roth_Arithmetic_Progressions.html, Formal proof development.

[12] Chelsea Edmonds, Angeliki Koutsoukou-Argyraki, and Lawrence C. Paulson. 2021. Szemerédi's Regularity Lemma. *Archive of Formal Proofs* (November 2021). https://isa-afp.org/entries/Szemeredi_Regularity.html, Formal proof development.

[13] Chelsea Edmonds, Angeliki Koutsoukou-Argyraki, and Lawrence C. Paulson. 2022. Formalising Szemerédi's Regularity Lemma and Roth's Theorem on Arithmetic Progressions in Isabelle/HOL. https://doi.org/10.48550/ARXIV.2207.07499

[14] Chelsea Edmonds and Lawrence C. Paulson. 2021. A Modular First Formalisation of Combinatorial Design Theory. In *Intelligent Computer Mathematics* (Timisoara, Romania), Fairouz Kamareddine and Claudio Sacerdoti Coen (Eds.). Springer-Verlag, Berlin, Heidelberg, 3–18. https://doi.org/10.1007/978-3-030-81097-9_1

[15] Jordan Ellenberg and Dion Gijswijt. 2017. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Annals of Mathematics* 185, 1 (2017), 339 – 343. https://doi.org/10.4007/annals.2017.185.1.8

[16] William Timothy Gowers. 2001. A New Proof of Szemerédi's Theorem. *Geometric & Functional Analysis GAFA* 11, 3 (2001), 465–588. https://doi.org/10.1007/s00039-001-0332-9

[17] William Timothy Gowers. 2022. Introduction to Additive Combinatorics. Lecture notes for Part III of the Mathematics Tripos taught at the University of Cambridge, available at https://drive.google.com/file/d/1ut0mUqSyPMweoxoDTfhXverEONyFgcuO/view.

[18] Lars Hupel. 2014. Properties of Random Graphs – Subgraph Containment. *Archive of Formal Proofs* (February 2014). https://isa-afp.org/entries/Random_Graph_Subgraph_Threshold.html, Formal proof development.

[19] Angeliki Koutsoukou-Argyraki, Mantas Bakšys, and Chelsea Edmonds. 2022. The Balog–Szemerédi–Gowers Theorem. *Archive of Formal Proofs* (November 2022). https://isa-afp.org/entries/Balog_Szemeredi_Gowers.html, Formal proof development.

[20] Angeliki Koutsoukou-Argyraki and Lawrence C. Paulson. 2022. Khovanskii's Theorem. *Archive of Formal Proofs* (September 2022). https://isa-afp.org/entries/Khovanskii_Theorem.html, Formal proof development.

[21] Angeliki Koutsoukou-Argyraki and Lawrence C. Paulson. 2022. The Plünnecke-Ruzsa Inequality. *Archive of Formal Proofs* (May 2022). https://isa-afp.org/entries/Pluennecke_Ruzsa_Inequality.html, Formal proof development.

[22] Katharina Kreuzer and Manuel Eberl. 2021. Van der Waerden's Theorem. *Archive of Formal Proofs* (June 2021). https://isa-afp.org/entries/Van_der_Waerden.html, Formal proof development.

[23] Bhavik Mehta. 2022. Formalising the Kruskal-Katona Theorem in Lean. In *Intelligent Computer Mathematics: 15th International Conference, CICM 2022, Tbilisi, Georgia, September 19–23, 2022, Proceedings* (Tbilisi, Georgia). Springer-Verlag, Berlin, Heidelberg, 75–91. https://doi.org/10.1007/978-3-031-16681-5_5

[24] Tobias Nipkow, Lawrence C. Paulson, and Markus Wenzel. 2002. *Isabelle/HOL: A Proof Assistant for Higher-Order Logic*. Springer. Online at http://isabelle.in.tum.de/dist/Isabelle/doc/tutorial.pdf.

[25] Benedikt Nordhoff and Peter Lammich. 2012. Dijkstra's Shortest Path Algorithm. *Archive of Formal Proofs* (January 2012). https://isa-afp.org/entries/Dijkstra_Shortest_Path.html, Formal proof development.

[26] Lars Noschinski. 2012. A Probabilistic Proof of the Girth-Chromatic Number Theorem. *Archive of Formal Proofs* (February 2012). https://isa-afp.org/entries/Girth_Chromatic.html, Formal proof development.

[27] Lars Noschinski. 2012. Proof Pearl: A Probabilistic Proof for the Girth-Chromatic Number Theorem. In *Interactive Theorem Proving. ITP 2012. (Lecture Notes in Computer Science, Vol. 7406)*, Lennart Beringer and Amy Felty (Eds.). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-32347-8_27

[28] Lars Noschinski. 2015. A Graph Library for Isabelle. *Mathematics in Computer Science* 9, 1 (March 2015), 23–39. https://doi.org/10.1007/s11786-014-0183-z

[29] Lawrence C. Paulson. 1986. Natural Deduction as Higher-Order Resolution. *J. Log. Program.* 3, 3 (October 1986), 237–258. https://doi.org/10.1016/0743-1066(86)90015-4

[30] Lawrence C. Paulson. 1989. The Foundation of a Generic Theorem Prover. *Journal of Automated Reasoning* 5, 3 (September 1989), 363–397. https://doi.org/10.1007/BF00248324

[31] Lawrence C. Paulson and Jasmin Christian Blanchette. 2012. Three years of experience with Sledgehammer, a Practical Link Between Automatic and Interactive Theorem Provers. In *IWIL 2010. The 8th International Workshop on the Implementation of Logics (EPiC Series in Computing, Vol. 2)*, Geoff Sutcliffe, Stephan Schulz, and Eugenia Ternovska (Eds.). EasyChair, 1–11. https://doi.org/10.29007/36dt

[32] Imre Z. Ruzsa. 1978. On the cardinality of $A + A$ and $A - A$. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976)*, Vol. 2. 933–938.

[33] Imre Z. Ruzsa. 2008. Sumsets and structure. Lecture notes, Institute of Mathematics, Budapest, available at https://www.math.cmu.edu/users/af1p/Teaching/AdditiveCombinatorics/Additive-Combinatorics.pdf.

[34] Yiannos Stathopoulos, Angeliki Koutsoukou-Argyraki, and Lawrence C. Paulson. 2020. Developing a Concept-Oriented Search Engine for Isabelle Based on Natural Language : Technical Challenges. In *5th Conference on Artificial Intelligence and Theorem Proving (AITP 2020), Aussois, France.* http://aitp-conference.org/2020/abstract/paper_9.pdf Informal proceedings.

[35] Yiannos Stathopoulos, Angeliki Koutsoukou-Argyraki, and Lawrence C. Paulson. 2020. SErAPIS: A Concept-Oriented Search Engine for the Isabelle Libraries Based on Natural Language. In *Isabelle Workshop 2020 (in virtual space).* https://files.sketis.net/Isabelle_Workshop_2020/Isabelle_2020_paper_4.pdf Informal proceedings.

[36] B. Sudakov, E. Szemerédi, and V. H. Vu. 2005. On a question of Erdős and Moser. *Duke Mathematical Journal* 129, 1 (2005), 129 – 155. https://doi.org/10.1215/S0012-7094-04-12915-X

[37] Ujkan Sulejmani, Manuel Eberl, and Katharina Kreuzer. 2022. The Hales–Jewett Theorem. *Archive of Formal Proofs* (September 2022). https://isa-afp.org/entries/Hales_Jewett.html, Formal proof development.

[38] Terence Tao and Van H. Vu. 2006. *Additive Combinatorics.* Cambridge University Press. https://doi.org/10.1017/CBO9780511755149

[39] René Thiemann. 2021. The Sunflower Lemma of Erdős and Rado. *Archive of Formal Proofs* (February 2021). https://isa-afp.org/entries/Sunflowers.html, Formal proof development.

[40] Markus Wenzel. 2002. *Isabelle, Isar - a Versatile Environment for Human Readable Formal Proof Documents.* PhD Thesis. Technical University Munich, Germany.

[41] Yufei Zhao. 2022. Graph Theory and Additive Combinatorics. Online at https://yufeizhao.com/gtacbook/. book draft.

Received 2022-09-21; accepted 2022-11-21