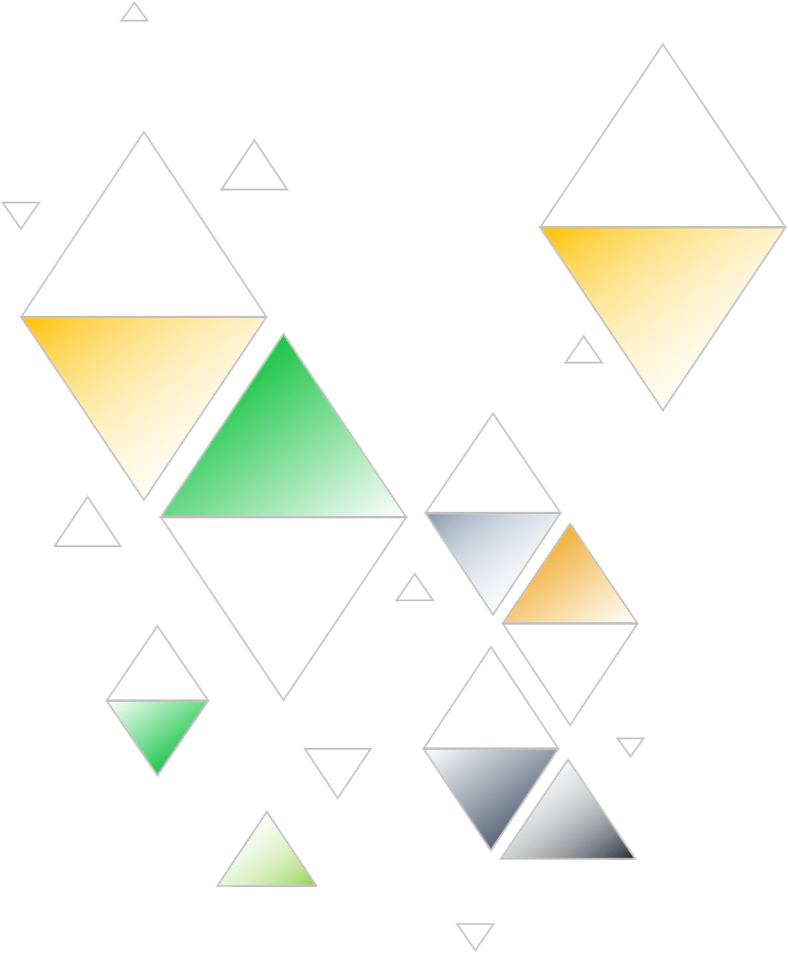
STRUCTURE IN PRIME GAPS - FORMALIZED

PROJECT PROPOSAL & BLUEPRINT

****

prepared for   
https://leanprover.zulipchat.com/

prepared by   
KAJANI KAUNDA

# 

**Proposed Channel Name: SPG-Formalized (SPG-F)**

**Introduction**

Welcome to the channel!

In this project, we aim to formalize the results presented in the paper [*Structure in Prime Gaps*](https://www.researchsquare.com/article/rs-4058806/latest) using the LEAN proof assistant. By leveraging the capabilities of LEAN4, we seek to ensure the correctness and robustness of the findings related to the existence of infinitely many pairs of prime numbers with specific gaps. This formalization will provide a rigorous verified foundation for the results and contribute to the broader effort of formalizing mathematics.

The paper [*Structure in Prime Gaps*](https://www.researchsquare.com/article/rs-4058806/latest) presents two main results the first of which is the claim that there exist structured gaps between primes and the second result is basically a corollary or special case of the first. These are stated as follows:

**Theorem 1**: For every prime pα, there exists infinitely many pairs of primes, (*p*n, *p*n+m), such that (*p*n+m − *p*n) = *p*α − 3, where n, α ≥ 3, m ≥ 1, and *p*n is the *n*th prime.

**Theorem 2**: There exist infinitely many pairs of primes with a gap of 2.

**Goal**

Our goal is to collaboratively formalize the proofs of these results using LEAN4.

**Definitions and Results**

We aim to formalize the following definitions and results.

**Definition 1**: Define a model of prime gaps as the Cayley Table *T* constructed from a *Cummutative Partial Groupoid* (*J*, +) based on a subset *J* of *Z* containing *all* prime numbers and their additive inverses, such that the elements of the first row of *T* are the primes and the elements of the first column of *T* are their additive inverses.

**NOTE**:

Definition 1 defines the infinite structure *T* which we use to model prime gaps.

**Definition 2**: Define a *v* x *w* sub-array *T*i of *T* such that *v, w* ≥ 2.

**NOTE**:

Definition 2 defines a sub-array which is later used to algebraically construct a pattern.

**Definition 3**: Define a 4-tuple *β* = (*A, B, L, E*) such that the values of its elements are the vertices of *T*i.

**NOTE**:

Definition 3 defines a tuple that is later used in analysis.

**Lemma 4.1**: Prove that: *A* + *E* = *B* + *L*.

**NOTE**:

We use this result in the subsequent proofs.

**Lemma 4.2**: Prove that: every prime number greater than 3 can be expressed in the form 6*n* ± 1.

(A known result).

**NOTE**:

We use this result in the subsequent proofs.

**Lemma 4.3**: Let *t*m,n be a term in *T* where the indexes *m* and *n* are zero based and refer to the rows and columns of *T* respectively.

Prove that:

For every prime *p*α ≥ 5, there exists an array *T*i ∈ *T* such that the following properties are simultaneously true;

Property 1 : *T*i.*A* + 3 ∈ {6*n* ± 1|*n* ∈ *N*1}

Property 2 : (*T*i.*B* + 3) − *T*i.*E* ∈ {6*n* ± 1|*n* ∈ *N*1}

Property 3 : *T*i.*L* ≡ 0 (mod 6)

Property 4 : *T*i.*A* = *T*i.*E*

Property 5 : *T*i.*B* + 3 ∈ {6*n* ± 1|*n* ∈ *N*1}

If and only if

for *T*i.*A* + 3 ∈ {6*n* − 1|*n* ∈ *N*1};

*T*i.*A* = 6*x* + 6*y* − 4

*T*i.*B* = 6*x* + 12*y* − 8

*T*i.*L* = 6*x*

*T*i.*E* = 6*x* + 6*y* − 4

for *T*i.*A* + 3 ∈ {6*n* + 1|*n* ∈ *N*1};

*T*i.*A* = 6*x* + 6*y* − 2

*T*i.*B* = 6*x* + 12*y* − 4

*T*i.*L* = 6*x*

*T*i.*E* = 6*x* + 6*y* − 2

where *n* ∈ *N*1, *x* < *n*, *y* > 0, *n* = *x* + *y*, (*T*i.*A* + 3) = *p*α and *T*i.*A* = *t*2,k.

**NOTE**:

For every prime *p*α ≥ 5, we algebraically prove the existence of a pattern using the sub-array *T*i, that defines a pair of integers *Q*i and *R*i such that *R*i − *Q*i = *p*α − 3.

**Lemma 4.4**: Let any array *T*i that satisfies Lemma 4.3 be referred to as a *Prime Array*.

Prove that: For every prime *p*α ≥ 5, there are infinitely many *Prime Arrays*.

**NOTE**:

This will show that for any prime *p*α ≥ 5 the pattern defined by the *Prime Array T*i occurs infinitely often and that consequently, the integers *Q*i and *R*i also occur infinitely often.

**Lemma 4.5**: Prove that: For every prime *p*α ≥ 5, there are infinitely many *Prime Arrays* and (*T*i.*B* + 3) and ((*T*i.*B* + 3) −*T*i.*E*) are prime.

**NOTE**:

This will show that for any prime *p*α ≥ 5, the *Prime Array* *T*i occurs infinitely often and *Q*i and *R*i are prime. Notice that here, (*T*i.*B* + 3) = *R*i and *Q*i = ((*T*i.*B* + 3) − *T*i.*E*).

**Theorem 1**: Prove that: For every prime *p*α, there exists infinitely many pairs of primes, (*p*n, *p*n+m), such that (*p*n+m − *p*n) = *p*α − 3, where n, α ≥ 3, m ≥ 1, and *p*n is the nth prime.

**NOTE**:

This result is implied from the previous results as demonstrated in the following steps:

**Step 1**: By **Lemma 4.5**, and the construction of *T*, *A* and *E* are prime gaps.

**Step 2**: By **Lemma 4.3**, *A* = *E*.

This is equivalent to: (*T*i.*A* + 3) − 3 = ((*T*i.*B* + 3) − ((*T*i.*B* + 3) − *T*i.*E*)).

**Step 3**: By **Lemma 4.5**, the following are prime; (*T*i.*A* + 3), ((*T*i.*B* + 3) − *T*i.*E*), (*T*i.*B* + 3).

**Step 4**: By **Lemma 4.5**, there are infinitely many pairs of primes defined as; (((*T*i.*B* + 3) − *T*i.*E*), (*T*i.*B* + 3)).

We can then see that the following statement is implied:

For every prime (*T*i.*A* + 3) ≥ 5, there exists infinitely many pairs of primes, (((*T*i.*B* + 3) − *T*i.*E*), (*T*i.*B* + 3)), such that ((*T*i.*B* + 3) − ((*T*i.*B* + 3) − *T*i.*E*)) = (*T*i.*A* + 3) − 3.

**Theorem 2**: This Result is just a special case of **Theorem 1**. I am sure this would be resolved by LEAN4 using the "refl" tactic.

**Community Etiquette**

* Be respectful and supportive.
* Ask questions if you need clarification.
* Share your knowledge and help.

**Timeline, Milestones and Deliverables**

There is no specifically set timeline for completion. Given the voluntary nature of such collaboration efforts, we can only enjoy the process, be thankful and hope that it takes as short a time as possible.

As for milestones, the ZULIP LEAN channel would be the best place to report and share results. We could setup the project on Github as well.

**Summary of Deliverables**

* Blueprint: We could build it from this document and the wonderful work by Patrick Massot.
* LEAN4 code: The Blueprint and the Github project could serve this purpose.
* Website: We could use the free Github pages facility.
* ArXiv e-print Archive Document: A tex document could be written.

**Scope - How You Can Contribute**

Formalization:

* Contributions from community members, are guided by the objectives and scope outlined in this proposal.
* Contribute by writing, translating, or verifying LEAN4 code to formalize the results.
* Help identify and fix any errors or inconsistencies in the formalization process/documentation.

Discussion:

* Participate in discussions to clarify concepts, share insights, and provide constructive feedback.
* Propose alternative approaches or solutions regarding the LEAN code to enhance the formalization.
* Regular updates and discussions on the Zulip LEAN chat channel.

Documentation:

* Assist in documenting the formalization process and results for better understanding and future reference.
* Review and refinement of Lean code to maintain accuracy and coherence with the original paper.

Outreach:

* Spread the word about our project to attract more contributors and experts.
* Engage with other communities and forums to gather more insights and feedback on the formalization process.

**Outcome and Conclusion**

A comprehensive and verified formalization of the proofs in [*Structure in Prime Gaps*](https://www.researchsquare.com/article/rs-4058806/latest), contributes to the broader effort of formalizing mathematical research using LEAN. We seek to conclude this effort with a paper published on the [arXiv.org e-Print archive](https://arxiv.org/) and possibly some suitable journal.

I am confident this effort like many others will go a long way towards establishing the teaching, formalization of mathematics and proof assistants in academia.

**Thank You**

Thank you in advance for your contributions and efforts in making this collaborative project a success.

Feel free to reach out with any queries, questions, or feedback on the paper. Let's work together to advance our understanding and formalization of prime pairs in mathematics!

# This documents VERSION HISTORY

| VERSION | REVISION DATE | DESCRIPTION OF CHANGES | REVISED BY |
| --- | --- | --- | --- |
| 1.0 | August 2024 | NONE | KAJANI KAUNDA |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

# PROPOSAL DETAILS

|  |  |  |  |
| --- | --- | --- | --- |
| **PROJECT TITLE** | SPG-FORMALIZED / SPG-F | | |
| **PROJECTED COMPLETION DATE** |  | **DATE OF PROPOSAL** | August 2024 |
| **PROJECT OWNER** | HTTPS://LEANPROVER.ZULIPCHAT.COM/ | **PROJECTED START DATE** | August 2024 |
| **ENDORSED BY** |  | **ENDORSED BY** |  |
| **ENDORSED BY** |  | **ENDORSED BY** |  |

|  |  |
| --- | --- |
| PREPARED BY | DATE |
| KAJANI KAUNDA | August 2024 | |

|  |  |
| --- | --- |
| UPDATED BY | DATE |
|  |  | |
|  |  | |

More rows/columns in the above tables may be added (or removed!) as necessary.



HURRAY!, to the ZULIPCHAT LEAN Community.

The community may be interested to see a visual representation of the results which are *self-evident* in the structure *T*. These would be irrefutable once formalized using LEAN4.

Consider Table 1, which represents the Cayley Table *T* from the algebraic structure (*J*, +). It is not immediately apparent if any useful pattern can be discerned from it. However, with the highlights in Table 2, a compelling *pattern* emerges, one that leads directly to **Theorem 1** from which **Theorem 2** is implied as seen in Table 3.

* Each pattern is defined and identified by the 4-tuple *β =* (*A, B, L, E*) formed from the elements in the vertices of *T*i. In Table 2, the first 4-tuple *β =* (*A, B, L, E*) for prime 23 is (20, 28, 12, 20).
* Every *Prime Array* *T*i, defines two pairs of primes: in Table 2, the *first pair* is (3, *p*α) or (3, 23) and the *second pair* is (((*B* + 3) + 0 – *E*), (*B* + 3)) or (11, 31).
* The *first pair* remains constant for all patterns related to any prime *p*α ≥ 5. Subsequent *second pairs* are unique for each *Prime Array* *T*i.
* *L* is always congruent to 0 mod 6.

**Legend for Tables:**

* **Table 1:** Represents the Cayley Table *T* of the *Commutative Partial Groupoid* structure (*J*, +) without immediately discernible patterns.
* **Table 2:** Highlights patterns where the 4-tuple *β* = (*A, B, L, E*) defines each pattern, with specific examples provided. The “Pattern” here is more elaborately defined in the sense that it consists of integers other than just *Q*i and *R*i.
* **Table 3:** Demonstrates the implication of results derived from the patterns observed in Table 2.
* **NOTE:** Remember that the Cayley Table *T* is a partial representation of an otherwise infinite structure. Therefore no generalization of these results to some “complete/larger” set is required.

**Remark:** These patterns are so fascinating that their formalization is inevitable!



**LEGEND**





