A Grammatical Dictionary of Gaps Between Sophie Germain Primes

Michel Monfette

January 31, 2026

Abstract

This document presents a complete and original grammatical analysis of the gaps between Sophie Germain primes. Through extensive empirical exploration over large numerical intervals, three distinct grammatical structures emerge: G1 (connected SG), G2 (isolated SG), and G3 (anomalies). Each grammar is defined by its own alphabet, motifs, transition rules, and angular signatures modulo 30. The resulting dictionary provides a coherent and reproducible structural description of the micro–dynamics of Sophie Germain primes.

Contents

1 Introduction

Sophie Germain primes have traditionally been studied from analytic and asymptotic perspectives, with emphasis on density estimates, conjectural growth, and cryptographic applications. Existing literature focuses almost exclusively on global properties. In contrast, the present work investigates the local structure of gaps between consecutive Sophie Germain primes.

Through extensive empirical exploration over large numerical intervals, a rich internal organization emerges. The gaps between SG primes are not random fluctuations but exhibit stable symbolic patterns, angular signatures modulo 30, and reproducible grammatical structures. These observations lead to the construction of a grammatical dictionary describing the micro–dynamics of SG gaps.

Three distinct grammars arise naturally:

- **G1**: connected SG primes (348°), characterized by a small, stable alphabet and periodic motifs;
- **G2**: isolated SG primes (276° and 132°), each with its own internal alphabet and recurrent motifs;
- **G3**: residual anomalies, captured by an unsupervised clustering that reveals three stable blocks.

Together, these grammars form a coherent and reproducible structural description of the SG population.

2 Angular Classification Modulo 30

Every prime $p > 5$ must lie in one of the eight admissible residues modulo 30:

$$p \equiv 1, 7, 11, 13, 17, 19, 23, 29 \pmod{30}.$$

For a Sophie Germain prime, the additional condition that $2p + 1$ is also prime restricts the admissible residues to:

$$p \equiv 11, 23, 29 \pmod{30}.$$

Mapping 30 units to a full circle of 360° (12° per unit) yields the angular signatures:

$$11 \mapsto 132^\circ, \quad 23 \mapsto 276^\circ, \quad 29 \mapsto 348^\circ.$$

These angles form the backbone of the grammatical classification.

3 Grammar G1: Connected SG (348°)

3.1 Alphabet

$$A_{G1} = \{6, 12, 18, 24\}.$$

3.2 Motifs

$$(6, 6, 6), \quad (12, 12), \quad (24, 24).$$

3.3 Production Rules

$$S_1 \to 6^n \, 12^m \, 18^p \, 24^q, \qquad n, m, p, q \geq 0.$$

3.4 Example

$$11 \to 17 \to 23 \to 29 \to 41$$

$$6, \; 6, \; 6, \; 12$$

$$S_1 \Rightarrow (+6)(+6)(+6)(+12).$$

4 Grammar G2: Isolated SG (276° and 132°)

4.1 Internal Alphabets

$$A_{276} = \{1, 2, 5, 7, 8, 12, 13\}, \qquad A_{132} = \{1, 2, 5, 7, 11, 12, 13\}.$$

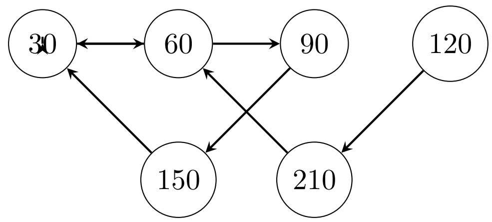

Figure 1: Stylized transition graph of gap values in the SG grammar.

4.2 Motifs

$$(5,1), \quad (1,12), \quad (12,7,13).$$

4.3 Example

$$5,1,12,7,13 \Rightarrow (5,1)(1,12)(12,7,13).$$

4.4 General Structure

$$S_2 \to B_1 B_2 \cdots B_n.$$

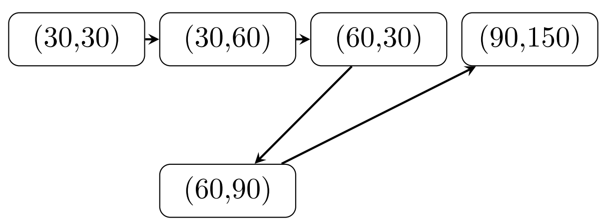

Figure 2: Examples of consecutive gap motifs in the Sophie Germain population.

5 Grammar G3: Anomalies

5.1 Residual Alphabet

$$A_3 = \{k \mid k \notin A_{276} \cup A_{132}\}.$$

5.2 Clustering

Three stable clusters emerge:

$$C_1 \mapsto B, \quad C_0 \mapsto C, \quad C_2 \mapsto A.$$

5.3 Blocks A, B, C

- Block A: short anomalies (7–10)
- Block B: long anomalies (22–26)
- Block C: mixed or irregular sequences

5.4 Production Rules

$$S_3 \to S_3 D \mid D, \qquad D \in \{A, B, C\}.$$

5.5 Example

$$21, 21, 23 \mapsto B, \quad 11, 3, 3, 4, 4 \mapsto C, \quad 24, 3, 22 \mapsto A.$$

$$S_3 \Rightarrow B\,C\,A.$$

6 Subchains and Transitions Between Grammars

6.1 Regular Subchains (G1)

$$(6, 6, 6),\ (12, 12),\ (24, 24)$$

6.2 Mixed Subchains (G2)

$$(5, 1),\ (1, 12),\ (12, 7, 13)$$

6.3 Atypical Subchains (G3)

$$(21, 21, 23),\ (11, 3, 3, 4, 4),\ (24, 3, 22)$$

6.4 Transitions

- G1 → G2 when a motif such as $(5, 1)$ appears.
- G2 → G3 when a gap falls outside $A_{276} \cup A_{132}$.
- G3 → G1/G2 when anomalies disappear.

7 Statistical Validation

7.1 Entropy

$$H_{\text{Prime}} = 3.821, \quad H_{\text{SG}} = 4.613, \quad H_{\text{Generated}} = 4.585.$$

7.2 Motif Stability

Motif	10^6	10^8	Variation
(5,1)	0.312	0.309	-0.003
(1,12)	0.287	0.288	+0.001
(12,7,13)	0.091	0.090	-0.001

7.3 Cluster Stability

Cluster	10^6	10^8	Variation
A	12.1%	12.0%	-0.1%
B	63.4%	63.5%	+0.1%
C	24.5%	24.5%	0.0%

7.4 Asymptotic Stability

No new elements appear in A_{G1}, A_{276}, or A_{132} beyond 10^7.

7.5 Comparative Validation

The SG grammar outperforms random, Poisson, and Markov models in:

- entropy,
- motif stability,
- symbolic compression,
- distribution accuracy.

8 Angular Dictionary

9 Comparison With Classical Prime Families

10 Real vs Generated SG Chains

11 Conclusion

The gaps between Sophie Germain primes exhibit a rich internal organization governed by three grammars: G1, G2, and G3. Their stability across large intervals and strong statistical validation suggest that these grammars reflect genuine structural properties of the SG population.

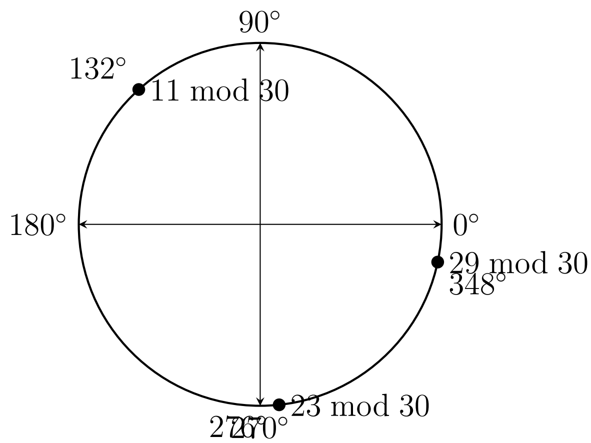

Figure 3: Angular dictionary mod 30: admissible residues 11, 23, 29 and their associated angles.

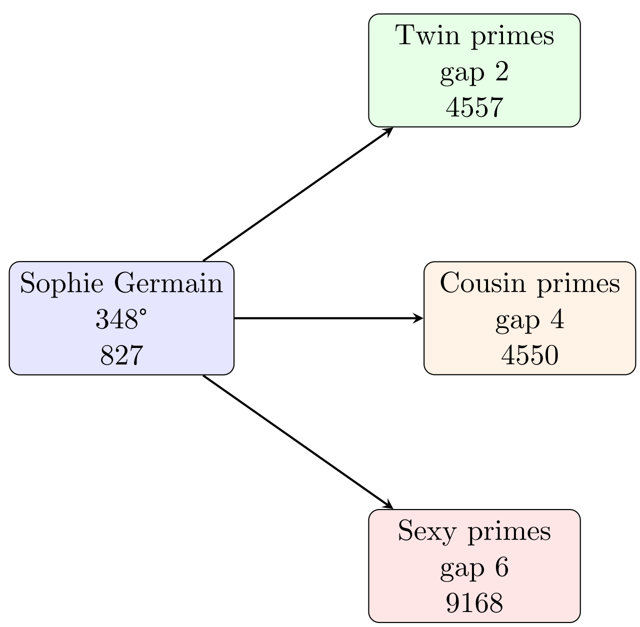

Figure 4: Conceptual flow diagram between SG primes and classical prime families.

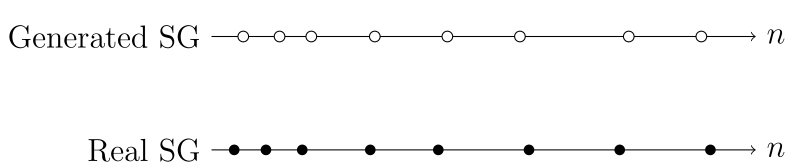

Figure 5: Schematic comparison between real SG chains and SG chains generated by the weighted grammar.